



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 104254

TO: Bao-Qun Li
Location: CM-1/7E07/8E12
Art Unit: 1648
Tuesday, September 23, 2003

Case Serial Number: 10/069883

From: Toby Port
Location: Biotech-Chem Library
CM1-6A04
Phone: 308-3534

toby.port@uspto.gov

Search Notes

Dear Examiner Li,

Here are the results of your search.
Please feel free to contact me if you have any questions.

Toby Port

THIS PAGE BLANK (USPTO)



STIC SEARCH RESULTS FEEDBACK FORM

Biotech-Chem Library

Questions about the scope or the results of the search? Contact *the searcher or contact*.

Mary Hale, Information Branch Supervisor
308-4258, CM1-1E01

Voluntary Results Feedback Form

➤ I am an examiner in Workgroup: Example: 1610

➤ Relevant prior art **found**, search results used as follows:

- ☐ 102 rejection
- ☐ 103 rejection
- ☐ Cited as being of interest.
- ☐ Helped examiner better understand the invention.
- ☐ Helped examiner better understand the state of the art in their technology.

Types of relevant prior art found:

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature
(journal articles, conference proceedings, new product announcements etc.)

➤ Relevant prior art **not found**:

- ☐ Results verified the lack of relevant prior art (helped determine patentability).
- ☐ Results were not useful in determining patentability or understanding the invention.

Comments:

Drop off or send completed forms to STIC/Biotech-Chem Library CM1 - Circ. Desk



THIS PAGE BLANK (USPTO)

STIC-Biotech/Ch mLib

104254

From: Li, Bao-Qun
Sent: Monday, September 22, 2003 3:21 PM
To: STIC-Biotech/ChemLib

Please go the sequence homology and interference search of SEQ ID NO: 1 of application SN. 10,069,883. Thanks.
AU 1648, 8E12.

RECEIVED
SEP 22 2003
STIC/CHEM. DIVISION
(STIC)

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: _____
Date Completed: _____
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:
NA Sequences: _____
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 23, 2003, 14:44:52 : Search time 85 Seconds

(without alignments)
1004.646 Million cell updates/sec

Title: US-10-069-883-1

Perfect score: 2871

Sequence: 1 MALPYHIFETVTLPSFTLL.....PPREISAQPLLRPNAGSS 538

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A.GeneSeq_19Jun03:*

1: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1980.DAT:*

2: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1981.DAT:*

3: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1982.DAT:*

4: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1983.DAT:*

5: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1984.DAT:*

6: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1985.DAT:*

7: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1986.DAT:*

8: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1987.DAT:*

9: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1988.DAT:*

10: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1989.DAT:*

11: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1990.DAT:*

12: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1991.DAT:*

13: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1992.DAT:*

14: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1993.DAT:*

15: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1994.DAT:*

16: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1995.DAT:*

17: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1996.DAT:*

18: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1997.DAT:*

19: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1998.DAT:*

20: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1999.DAT:*

21: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2000.DAT:*

22: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2001.DAT:*

23: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2002.DAT:*

24: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2871	100.0	538	22 AAB75138	HERV-W envelope pr
2	2871	100.0	538	22 AAB67652	Amino acid sequenc
3	2871	100.0	540	21 ABA40980	Human retroviral H
4	2871	100.0	685	21 ABA40988	Human retroviral e
5	2871	100.0	846	21 ABA41050	HERV-7g env protei
6	2868	99.9	538	20 AAY08622	Human secreted pro
7	2868	99.9	538	20 AAY67313	Human secreted pro
8	2868	99.9	538	23 AAE14540	Human syncytin pro
9	2791	97.2	538	20 AAW99345	Human endogenous r

10	2558	89.1	542	23 AAE25054	Human EMBRY-2 prot
11	2503.5	87.2	541	21 AAB08195	Amino acid sequenc
12	2490	86.7	542	21 AAB19069	Amino acid sequenc
13	2314	80.6	493	19 AAW71068	Multiple sclerosis
14	2314	80.6	493	20 AAW99553	Protein encoded by
15	1234.5	43.0	263	22 ABB51885	Human liver peptid
16	1234.5	43.0	263	22 ABB31812	Peptide #4463 enco
17	1234.5	43.0	263	22 ABB37043	Peptide #4549 enco
18	1234.5	43.0	263	22 ABB22358	Protein #4357 enco
19	1234.5	43.0	263	22 AAM57770	Human brain expres
20	1234.5	43.0	263	22 AAM70185	Human bone marrow
21	1234.5	43.0	263	22 AAM18008	Peptide #4442 enco
22	1234.5	43.0	263	22 AAM30515	Peptide #4556 enco
23	1234.5	43.0	263	22 AAM05648	Peptide #4330 enco
24	1234.5	43.0	263	23 ABB39819	Human peptid enco
25	1178.5	41.0	633	22 ABB28306	Novel human diagn
26	1174.5	40.9	368	22 ABB28311	Novel human diagn
27	1171	40.8	531	22 ABB63036	Human polypeptide
28	1126.5	39.2	446	22 ABB07532	Novel human diagn
29	1126.5	39.2	446	22 ABB20007	Novel human diagn
30	686	23.9	120	22 AAB75137	HERV-W truncated e
31	672	23.4	162	19 AAW71069	Multiple sclerosis
32	672	23.4	162	20 AAW99554	Protein encoded by
33	585.5	20.4	567	16 AAR71700	Spleen necrosis v1
34	563.5	19.6	584	22 ABB05606	Novel human diagn
35	563.5	19.6	584	22 ABB19851	Novel human diagn
36	563.5	19.6	584	22 ABB22871	Novel human diagn
37	554	19.3	586	22 ABB19838	Novel human diagn
38	554	19.3	1167	22 ABB05605	Novel human diagn
39	554	19.3	1186	22 ABB06819	Novel human diagn
40	554	19.3	1340	22 ABB06717	Novel human diagn
41	534.5	18.6	587	22 ABB12484	Novel human diagn
42	534.5	18.6	1131	22 ABB18584	Novel human diagn
43	530.5	18.5	469	23 ABB03726	Human ovary specif
44	506	17.6	538	23 ABB69150	Human polypeptid
45	488	17.0	231	22 ABB12209	Novel human diagn

ALIGNMENTS

RESULT 1	
AAB75138	
ID	AAB75138 standard; protein; 538 AA.
XX	
AC	AAB75138;
XX	
DT	08-AUG-2001 (first entry)
XX	
DE	HERV-W envelope protein G.
XX	
KW	Human endogenous retrovirus; HERV-W; HERV; chromosome 7; env protein;
KW	envelope protein; multiple sclerosis-related superantigen; vaccine;
KW	surface antigen; transmembrane; multiple sclerosis; neutroprotective;
KW	antisense-therapy; autoimmune disorder.
XX	
OS	Human endogenous retrovirus.
XX	
FH	Key
FT	271..317
FT	Region
FT	/label= SU
FT	/note= "surface protein"
FT	318..538
FT	/label= TM
FT	/note= "transmembrane subunit"
FT	317..318
FT	/note= "SU-TM putative cleavage site"
XX	
XX	Cleavage-site
XX	
XX	WO200131021-A1.
XX	
XX	03-MAY-2001.
XX	
XX	30-OCT-2000; 2000MO-EP10659.

```

XX 28-OCT-1999; 99EP-0402690.
PR (UYCE-) UNIV GENEVE.
XX
XX Conrad B, Mach B;
XX WPI: 2001-316336/33.
XX DR N-PSDB: AAH20070.
XX
PR New human retrovirus HERV-W ENV proteins/peptides having superantigen
PT activity useful for diagnosing and treating multiple sclerosis -
XX
XX Claim 1; Fig 7; 94pp: English.
XX
CC On the basis of the PBS t-RNA motif used for the classification of human
CC endogenous retrovirus (HERVs) the full length endogenous provirus which
CC was located on the long arm of human chromosome 7 (7q21-22) has been
CC designated HERV-W. The present invention describes proteins or peptides
CC (I) having superantigen (SAG) activity comprising the ENV protein (ENV)
CC of HERV-W, the surface protein (SU) and transmembrane (TM) sub-units. (I)
CC have neuroprotective activity, and can be used in: vaccines; antisense-
CC therapy; and HERV-W SAG activity-inhibitors. (I) and encoding DNA/RNA are
CC useful for diagnosing multiple sclerosis (MS) or HERV-W-associated
CC disorders. (I) are also useful for identifying substances (and optionally
CC recovering) capable of binding to a retroviral superantigen associated
CC with MS, substances capable of blocking SAG activity and substances
CC capable of blocking transcription or translation of HERV-W retroviral
CC superantigen. A protein or peptide derived from (I), modified to be
CC devoid of SAG activity and being capable of generating an immune response
CC against HERV-W retroviral SAG is useful in therapy. Nucleic acid
CC molecules encoding (I) are useful as vaccines against MS. Substances
CC capable of blocking SAG activity, capable of binding to a retroviral
CC superantigen associated with MS, or capable of blocking transcription or
CC translation of HERV-W retroviral superantigen for use in treating or
CC preventing MS, obtained using (I) are useful for the treatment and
CC prevention of MS. (I) and nucleic acids encoding them are useful for
CC diagnosing autoimmune disease. The present sequence represents the
CC specifically claimed envelope protein of HERV-W designated G.
XX
XX Sequence 538 AA:
SO
Query Match 100.0%; Score 2871; DB 22; Length 538;
Best Local Similarity 100.0%; Pred. No. 3.1e-259;
Matches 538; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALPYHIFLFTVLLPSFTLTAPPCRCMTSSPYOEFLMRQRCGNIDAPSRSLSKGTP 60
DB 1 MALPYHIFLFTVLLPSFTLTAPPCRCMTSSPYOEFLMRQRCGNIDAPSRSLSKGTP 60
QY 61 TFTAHTHPRNCYSATLCMHANTHYWTGKMINPSCPGGLGVYVCWTFYTOTGMSDGGV 120
DB 61 TFTAHTHPRNCYSATLCMHANTHYWTGKMINPSCPGGLGVYVCWTFYTOTGMSDGGV 120
QY 121 QDOAREKHVKVVISQLTRVHGCTSSPYKGLDLSKILHETRLTRTLVSLFNTTLGLAEVSA 180
DB 121 QDOAREKHVKVVISQLTRVHGCTSSPYKGLDLSKILHETRLTRTLVSLFNTTLGLAEVSA 180
QY 121 QDOAREKHVKVVISQLTRVHGCTSSPYKGLDLSKILHETRLTRTLVSLFNTTLGLAEVSA 180
DB 121 QDOAREKHVKVVISQLTRVHGCTSSPYKGLDLSKILHETRLTRTLVSLFNTTLGLAEVSA 180
QY 181 QNPNTCWTCLPLNRPVYSIPVPPQMNFSFEINTVSVLVPVYSNLEITHTSNLTCKVF 240
DB 181 QNPNTCWTCLPLNRPVYSIPVPPQMNFSFEINTVSVLVPVYSNLEITHTSNLTCKVF 240
QY 241 SNTTYTTSOCIRMTPTPTQIVCLPSGIFVCGTSAYNCLGSSSESMCFSLFLVPPMTIY 300
DB 241 SNTTYTTSOCIRMTPTPTQIVCLPSGIFVCGTSAYNCLGSSSESMCFSLFLVPPMTIY 300
QY 301 TEODLYSVIVISKPRKRVPILPFIYAGAVLGLGIGIGITTSOIFYKLSQELNGDMER 360
DB 301 TEODLYSVIVISKPRKRVPILPFIYAGAVLGLGIGIGITTSOIFYKLSQELNGDMER 360
QY 361 VADSLVLTQDQNLNSLAANVLYONRRALDLTFAERGCTCLFLGEECCYYVNOGIYTEKYE 420
DB 361 VADSLVLTQDQNLNSLAANVLYONRRALDLTFAERGCTCLFLGEECCYYVNOGIYTEKYE 420

```

```

QY 421 IRDRIORRAEELRNTGPMGLISQMPWILPFLGPLAAITILLFPGCIENLLVNFVSSRI 480
DB 421 IRDRIORRAEELRNTGPMGLISQMPWILPFLGPLAAITILLFPGCIENLLVNFVSSRI 480
QY 481 EAVKLQMBPKMQSKTKIYRRPLDRPASPRSDVNDIKGTPPEETISAQPLRNSAGSS 538
DB 481 EAVKLQMBPKMQSKTKIYRRPLDRPASPRSDVNDIKGTPPEETISAQPLRNSAGSS 538
RESULTS
AAB67652
ID AAB67652 standard; Protein; 538 AA.
XX
XX AAB67652:
XX
XX 29-MAY-2001 (first entry)
XX
XX Amino acid sequence of a human endogenous retrovirus envelope protein.
DE
XX Envelope protein; HERV; syncytia formation; placental development;
XX syncytia; cancer; cell adhesion.
XX
XX Human endogenous retrovirus.
OS
XX
XX WO200116171-A1.
XX
XX 08-MAR-2001.
XX
XX 01-SEP-2000; 2000WO-FR02429.
XX
XX 01-SEP-1999; 99EP-0011141.
XX
XX 15-SEP-1999; 99EP-0011793.
XX
XX (INNR ) BIO MERIEUX.
XX
XX (INNR ) INST NAT SANTE & RECH MEDICALE.
XX
XX Mallet F, Coasset F, Blond J, Lavillette D, Boulton O, Ruggieri A;
XX
XX WPI: 2001-226676/23.
XX
XX N-PSDB: AAF35630.
XX
XX Detecting expression of human endogenous retrovirus envelope protein in
XX cells of a tissue or culture, from its ability to induce syncytia -
XX
XX Claim 1; Page 41-44; 57pp; French.
XX
XX The present sequence represents a human endogenous retrovirus envelope
XX protein. The specification describes a method for detecting expression
XX of an envelope protein from a human endogenous retrovirus (HERV), in
XX cells, of a tissue or culture. The method comprises detecting syncytia
XX formation due to the fusogenic properties of the envelope protein.
XX Envelope polypeptides and polynucleotides are used to produce
XX therapeutic or prophylactic compositions, particularly for treatment of
XX cancer, to correct defects in placental development (or other natural
XX formation of other types of syncytia), and to promote adhesion of cells
XX in grafts or cellular repair processes. Expression of sequences
XX antisense to the polynucleotide are used to prevent formation of
XX syncytia.
XX
XX Sequence 538 AA:
SO
Query Match 100.0%; Score 2871; DB 22; Length 538;
Best Local Similarity 100.0%; Pred. No. 3.1e-259;
Matches 538; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALPYHIFLFTVLLPSFTLTAPPCRCMTSSPYOEFLMRQRCGNIDAPSRSLSKGTP 60
DB 1 MALPYHIFLFTVLLPSFTLTAPPCRCMTSSPYOEFLMRQRCGNIDAPSRSLSKGTP 60
QY 61 TFTAHTHPRNCYSATLCMHANTHYWTGKMINPSCPGGLGVYVCWTFYTOTGMSDGGV 120
DB 61 TFTAHTHPRNCYSATLCMHANTHYWTGKMINPSCPGGLGVYVCWTFYTOTGMSDGGV 120

```



```

QY 121 QDAREKHVKEVISQLTRVHGTSPPYKGLDLSKLTHTLRVLSLFTTLGLHEVSA 180
    |||
DQ 121 QDAREKHVKEVISQLTRVHGTSPPYKGLDLSKLTHTLRVLSLFTTLGLHEVSA 180
QY 181 QNPTNCWICLPLNFRPYVSIPEVEQNNFSTEINTTSVLVGPLVSNLEITHSNLCVKE 240
    |||
DQ 181 QNPTNCWICLPLNFRPYVSIPEVEQNNFSTEINTTSVLVGPLVSNLEITHSNLCVKE 240
QY 241 SNTTYTNSQCIHWVTPPTQIVCLPSGIFVCGTSAYRCLNGSSEKCFSLFVPPMTIY 300
    |||
DQ 241 SNTTYTNSQCIHWVTPPTQIVCLPSGIFVCGTSAYRCLNGSSEKCFSLFVPPMTIY 300
QY 301 TEODLXSVYSKPRNKRVPILPFVIGAGVIGALGTGTTSTOFYKLSOELNGDME 360
    |||
DQ 301 TEODLXSVYSKPRNKRVPILPFVIGAGVIGALGTGTTSTOFYKLSOELNGDME 360
QY 361 VADSLVTLQDQNSLAAYVLQNNRRLDLTAERGCTCLFGECCYVNOGSIYTERVKE 420
    |||
DQ 361 VADSLVTLQDQNSLAAYVLQNNRRLDLTAERGCTCLFGECCYVNOGSIYTERVKE 420
QY 421 IRDRIORRAEELNRTGPMGLLSQMPWILPFLGPIAAIILLFGPCIFNLVNFVSSRI 480
    |||
DQ 421 IRDRIORRAEELNRTGPMGLLSQMPWILPFLGPIAAIILLFGPCIFNLVNFVSSRI 480
QY 481 EAVKLQMEPRKMSKTKIYRRPLDRPASPRSDVNDIGTPPEEISAQPLLRPSAGSS 538
    |||
DQ 481 EAVKLQMEPRKMSKTKIYRRPLDRPASPRSDVNDIGTPPEEISAQPLLRPSAGSS 538

```

RESULT 3

ABP40980 standard; Protein; 540 AA.

```

ID ABP40980 standard; Protein; 540 AA.
XX
AC ABP40980;
XX
DT 01-AUG-2002. (first entry)
XX
DE Human retroviral HERV-7q env peptide #3.
XX
KW Autoimmune disease; HERV-7q; chromosome 7q; immunotherapy;
    multiple sclerosis.
XX
OS Human retrovirus.
XX
PN WO967395-A1.
XX
PD 29-DEC-1999.
XX
PF 23-JUN-1999; 99MO-FR01513.
XX
PR 23-JUN-1998; 98FR-0007920.
XX
PA (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
XX
PI Alliel PM, Perin J, Rieger F;
XX
DR WPI; 2000-160587/14.
XX
PT New nucleic acid sequences of human endogenous retrovirus, HERV-7q,
    used for diagnosis, treatment and prevention of autoimmune and
    neurological diseases.
XX
PS Claim 22; Fig 4; 225pp; French.
XX
CC The present invention relates to new nucleic acid sequences of human
    endogenous retrovirus, HERV-7q, which is located on chromosome 7q.
    Regulatory elements associated with HERV-7q may alter expression of other
    genes (even remote genes) on the same chromosome, inducing immunological
    and/or neurological changes (which may be pathological or protective/
    curative). HERV-7q peptides can be used to improve efficiency of the
    immune response, e.g. in immunotherapy. HERV-7q peptides and their coding
    sequences can be used in immunogenic or vaccinating compositions, for

```

CC protection against autoimmune diseases, particularly multiple sclerosis.
 CC The peptides may also be used (by sequence comparison) to detect/identify
 CC endogenous retroviruses that are abnormally expressed in cancer,
 CC neuropathologies or other autoimmune diseases. The present sequence was
 CC used to illustrate the invention.

Sequence 540 AA;

Query Match 100.0%; Score 2871; DB 21; Length 540;
 Best Local Similarity 100.0%; Pred. No. 3.1e-259;
 Matches 538; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MALPYHFLFTVLLPFTTLAPPCCOMSSSPYQEFELRMORGNIDAPSVLSKGP 60
    |||
DQ 3 MALPYHFLFTVLLPFTTLAPPCCOMSSSPYQEFELRMORGNIDAPSVLSKGP 62
QY 61 TPTAHTHMPRNCYHSATLCMHANTHYWGKIMIPSCPGIGVVCWTYFTQTMSDGGCV 120
    |||
DQ 63 TPTAHTHMPRNCYHSATLCMHANTHYWGKIMIPSCPGIGVVCWTYFTQTMSDGGCV 122
QY 121 QDAREKHVKEVISQLTRVHGTSPPYKGLDLSKLTHTLRVLSLFTTLGLHEVSA 180
    |||
DQ 123 QDAREKHVKEVISQLTRVHGTSPPYKGLDLSKLTHTLRVLSLFTTLGLHEVSA 182
QY 181 QNPTNCWICLPLNFRPYVSIPEVEQNNFSTEINTTSVLVGPLVSNLEITHSNLCVKE 240
    |||
DQ 183 QNPTNCWICLPLNFRPYVSIPEVEQNNFSTEINTTSVLVGPLVSNLEITHSNLCVKE 242
QY 241 SNTTYTNSQCIHWVTPPTQIVCLPSGIFVCGTSAYRCLNGSSEKCFSLFVPPMTIY 300
    |||
DQ 243 SNTTYTNSQCIHWVTPPTQIVCLPSGIFVCGTSAYRCLNGSSEKCFSLFVPPMTIY 302
QY 301 TEODLXSVYSKPRNKRVPILPFVIGAGVIGALGTGTTSTOFYKLSOELNGDME 360
    |||
DQ 303 TEODLXSVYSKPRNKRVPILPFVIGAGVIGALGTGTTSTOFYKLSOELNGDME 362
QY 361 VADSLVTLQDQNSLAAYVLQNNRRLDLTAERGCTCLFGECCYVNOGSIYTERVKE 420
    |||
DQ 363 VADSLVTLQDQNSLAAYVLQNNRRLDLTAERGCTCLFGECCYVNOGSIYTERVKE 422
QY 421 IRDRIORRAEELNRTGPMGLLSQMPWILPFLGPIAAIILLFGPCIFNLVNFVSSRI 480
    |||
DQ 423 IRDRIORRAEELNRTGPMGLLSQMPWILPFLGPIAAIILLFGPCIFNLVNFVSSRI 482
QY 481 EAVKLQMEPRKMSKTKIYRRPLDRPASPRSDVNDIGTPPEEISAQPLLRPSAGSS 538
    |||
DQ 483 EAVKLQMEPRKMSKTKIYRRPLDRPASPRSDVNDIGTPPEEISAQPLLRPSAGSS 540

```

RESULT 4

ABP40988 standard; Protein; 685 AA.

```

ID ABP40988 standard; Protein; 685 AA.
XX
AC ABP40988;
XX
DT 01-AUG-2002 (first entry)
XX
DE Human retroviral env ORF protein.
XX
KW Autoimmune disease; HERV-7q; chromosome 7q; immunotherapy;
    multiple sclerosis.
XX
OS Human retrovirus.
XX
PN WO967395-A1.
XX
PD 29-DEC-1999.
XX
PF 23-JUN-1999; 99MO-FR01513.
XX
PR 23-JUN-1998; 98FR-0007920.
XX
PA (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.

```

XX Alliel PM, Perin J, Rieger F;
PI WPI: 2000-160587/14.
XX
XX New nucleic acid sequences of human endogenous retrovirus, HERV-7q,
XX used for diagnosis, treatment and prevention of autoimmune and
XX neurological diseases -
XX
XX Claim 22: Page 152-154; 225pp; French.
XX
XX The present invention relates to new nucleic acid sequences of human
XX endogenous retrovirus, HERV-7q, which is located on chromosome 7q.
XX Regulatory elements associated with HERV-7q may alter expression of other
XX genes (even remote genes) on the same chromosome, inducing immunological
XX and/or neurological changes (which may be pathological or protective/
XX curative). HERV-7q peptides can be used to improve efficiency of the
XX immune response, e.g. in immunotherapy. HERV-7q peptides and their coding
XX sequences can be used in immunogenic or vaccinating compositions, for
XX protection against autoimmune diseases, particularly multiple sclerosis.
XX The peptides may also be used (by sequence comparison) to detect/identify
XX endogenous retroviruses that are abnormally expressed in cancer.
XX neuropathologies or other autoimmune diseases. The present sequence was
XX used to illustrate the invention.

SQ Sequence 685 AA;

Query Match 100.0%; Score 2871; DB 21; Length 685;
Best Local Similarity 100.0%; Pred. No. 4,4e-259;
Matches 538; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALPYHIFLFTVLPSPFTLTAPPCCRCMTSSPYOEFILMRQRPENIDAPSYRSLSKCTP 60
DB 131 MALPYHIFLFTVLPSPFTLTAPPCCRCMTSSPYOEFILMRQRPENIDAPSYRSLSKCTP 190
QY 61 TFTANTHPRNCYHSATLCMAHNTHYTGKMINPSCPGGLGVTVVCWTFYFTQMSDGGCV 120
DB 191 TFTANTHPRNCYHSATLCMAHNTHYTGKMINPSCPGGLGVTVVCWTFYFTQMSDGGCV 250
QY 121 ODOAREKHVKEVISOLTRVHGHTSSPYKGLDLSKIHETLRTHRLVLSLNTLTGLHEVSA 180
DB 251 ODOAREKHVKEVISOLTRVHGHTSSPYKGLDLSKIHETLRTHRLVLSLNTLTGLHEVSA 310
QY 181 ONPTNCWICLPLNFRPYVSIPEEQMNNFSTEINTSVLVGPLYNSLEITHTSNLTCKVF 240
DB 311 ONPTNCWICLPLNFRPYVSIPEEQMNNFSTEINTSVLVGPLYNSLEITHTSNLTCKVF 370
QY 241 SNTTYTNSOCIRWTPPTQIVCLPSGIFVCGTSAYRCLNGSSSEKCFLSFLVPPMTIY 300
DB 371 SNTTYTNSOCIRWTPPTQIVCLPSGIFVCGTSAYRCLNGSSSEKCFLSFLVPPMTIY 430
QY 301 TEODLYSVISKPRKRKRPILPFVIGAGVLGALGIGITTSQFYKLSOELNGDMER 360
DB 431 TEODLYSVISKPRKRKRPILPFVIGAGVLGALGIGITTSQFYKLSOELNGDMER 490
QY 361 VADSLVTLQDOLNSIAAVVLONNRALDLTLAERGCTCLFLGEECCYVYVNSQIVTEKYE 420
DB 491 VADSLVTLQDOLNSIAAVVLONNRALDLTLAERGCTCLFLGEECCYVYVNSQIVTEKYE 550
QY 421 IRDRIRORAEELRNTPGWLSSQMPWILPFLGPIAATITLLLFEPCCIFNLVNVSSRI 480
DB 551 IRDRIRORAEELRNTPGWLSSQMPWILPFLGPIAATITLLLFEPCCIFNLVNVSSRI 610
QY 481 EAVKLQMPKMSKTKIYRRPLDRPASPRSDVNDIKGTPPEBISAAQPLLRNSAGSS 538
DB 611 EAVKLQMPKMSKTKIYRRPLDRPASPRSDVNDIKGTPPEBISAAQPLLRNSAGSS 668

RESULT 5
ID ABP41050 standard: Protein; 846 AA.
XX
AC ABP41050;

XX 01-AUG-2002 (first entry)
XX
XX HERV-7q env protein #1.
XX
XX Autoimmune disease; HERV-7q; chromosome 7q; immunotherapy;
XX multiple sclerosis.
XX
XX Human retrovirus.
XX
XX WO967395-A1.
XX
XX 29-DEC-1999.
XX
XX 23-JUN-1999; 99WO-FR01513.
XX
XX 23-JUN-1998; 98FR-0007920.
XX
XX (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
XX
XX Alliel PM, Perin J, Rieger F;
XX WPI: 2000-160587/14.
XX
XX New nucleic acid sequences of human endogenous retrovirus, HERV-7q,
XX used for diagnosis, treatment and prevention of autoimmune and
XX neurological diseases -
XX
XX Example 4: Fig 19; 225pp; French.

The present invention relates to new nucleic acid sequences of human
endogenous retrovirus, HERV-7q, which is located on chromosome 7q.
Regulatory elements associated with HERV-7q may alter expression of other
genes (even remote genes) on the same chromosome, inducing immunological
and/or neurological changes (which may be pathological or protective/
curative). HERV-7q peptides can be used to improve efficiency of the
immune response, e.g. in immunotherapy. HERV-7q peptides and their coding
sequences can be used in immunogenic or vaccinating compositions, for
protection against autoimmune diseases, particularly multiple sclerosis.
The peptides may also be used (by sequence comparison) to detect/identify
endogenous retroviruses that are abnormally expressed in cancer.
neuropathologies or other autoimmune diseases. The present sequence was
used to illustrate the invention.

SQ Sequence 846 AA;

Query Match 100.0%; Score 2871; DB 21; Length 846;
Best Local Similarity 100.0%; Pred. No. 6,1e-259;
Matches 538; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALPYHIFLFTVLPSPFTLTAPPCCRCMTSSPYOEFILMRQRPENIDAPSYRSLSKCTP 60
DB 291 MALPYHIFLFTVLPSPFTLTAPPCCRCMTSSPYOEFILMRQRPENIDAPSYRSLSKCTP 350
QY 61 TFTANTHPRNCYHSATLCMAHNTHYTGKMINPSCPGGLGVTVVCWTFYFTQMSDGGCV 120
DB 351 TFTANTHPRNCYHSATLCMAHNTHYTGKMINPSCPGGLGVTVVCWTFYFTQMSDGGCV 410
QY 121 ODOAREKHVKEVISOLTRVHGHTSSPYKGLDLSKIHETLRTHRLVLSLNTLTGLHEVSA 180
DB 411 ODOAREKHVKEVISOLTRVHGHTSSPYKGLDLSKIHETLRTHRLVLSLNTLTGLHEVSA 470
QY 181 ONPTNCWICLPLNFRPYVSIPEEQMNNFSTEINTSVLVGPLYNSLEITHTSNLTCKVF 240
DB 471 ONPTNCWICLPLNFRPYVSIPEEQMNNFSTEINTSVLVGPLYNSLEITHTSNLTCKVF 530
QY 241 SNTTYTNSOCIRWTPPTQIVCLPSGIFVCGTSAYRCLNGSSSEKCFLSFLVPPMTIY 300
DB 531 SNTTYTNSOCIRWTPPTQIVCLPSGIFVCGTSAYRCLNGSSSEKCFLSFLVPPMTIY 590
QY 301 TEODLYSVISKPRKRKRPILPFVIGAGVLGALGIGITTSQFYKLSOELNGDMER 360
DB 591 TEODLYSVISKPRKRKRPILPFVIGAGVLGALGIGITTSQFYKLSOELNGDMER 650

```

OY 361 VADSLVTLDQDLSLAAYVQLQNRALDILLTAERGCTCLFGECCYYVQSGIVTEKYKE 420
    |||||||
DB 651 VADSLVTLDQDLSLAAYVQLQNRALDILLTAERGCTCLFGECCYYVQSGIVTEKYKE 710
OY 421 IRDRIQRAAEELRNTGPMGLLSQMPMPILPFLGPIAAIILLFGPCIFNLVNFVSSRI 480
    |||||||
DB 711 IRDRIQRAAEELRNTGPMGLLSQMPMPILPFLGPIAAIILLFGPCIFNLVNFVSSRI 770
OY 481 EAVKLOMEPRKMSKTKIYRRPLDRPASPSDVNDIKGTPPEEISAQPLLRPSAGSS 538
    |||||||
DB 771 EAVKLOMEPRKMSKTKIYRRPLDRPASPSDVNDIKGTPPEEISAQPLLRPSAGSS 828

RESULT 6
AA08622
ID AAY08622 standard; Protein; 538 AA.
AC AAY08622;
XX
XX
DT 10-AUG-1999 (first entry)
DE Human secreted protein AJ172_2.
XX
XX Secreted protein; testes; brain; blood; placenta; human; murine; thymus;
KW bone marrow; treatment; prevention; nutrition; cytokine; immune; vaccine;
KW cell proliferation; cell differentiation; suppressor; tumor inhibition;
KW haematopoiesis regulator; activin; inhibin; chemotactic; chemokine;
KW haemostatic; thrombolytic; receptor; ligand; anti-inflammatory; tumor;
KW cadherin; tumor invasion suppressor; gene therapy; tissue growth.
XX
XX Homo sapiens.
OS
XX
XX WO9926972-A1.
XX
XX 03-JUN-1999.
XX
XX 17-NOV-1998; 98WO-US24614.
XX
XX 20-OCT-1998; 98US-0175928.
XX
XX 21-NOV-1997; 97US-0976110.
XX
XX 18-MAY-1998; 98US-0080478.
XX
XX (GENE ) GENETICS INST INC.
XX
XX Collins-Racie LA, Evans C, Jacobs K, Lavallic ER.
XX
XX McCoy JM, Werberg D, Treacy M.
XX
XX WPI; 1999-357813/30.
XX
XX N-PSDB; AAX7526.
XX
XX New polynucleotides encoding secreted proteins
XX
XX Claim 14a; Page 101-103; 142pp; English.
XX
XX This invention describes novel human secreted proteins encoded by
XX polynucleotides isolated from human adult testes, adult brain, adult
XX blood or adult placenta, or murine adult bone marrow or thymus cDNA
XX libraries. The products of the invention are predicted to have biological
XX activities which would make them suitable for treating, preventing or
XX ameliorating medical conditions in humans and animals, although no
XX supporting data is given. Suggested activities include nutritional
XX activity, cytokine and cell proliferation/differentiation activity,
XX immune stimulating (e.g. as vaccines) or suppressing activity,
XX haematopoiesis regulating activity, tissue growth activity,
XX activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
XX and thrombolytic activity, receptor/ligand activity, anti-inflammatory
XX activity, cadherin/tumour invasion suppressor activity, and tumour
XX inhibition activity. The polynucleotides are also stated to be useful
XX for gene therapy.
XX
XX Sequence 538 AA;
SQ

```

```

Query Match          99.9%; Score 2868; DB 20; Length 538;
Best local Similarity 99.8%; Pred. No. 5; 9e-259;
Matches 537; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 MALPHYIFLFTVLLSEFTLTAPPCCGMTSSSFYQEFLLRMQRPQIDAPSTRSLSKGP 60
    |||||||
DB 1 MALPHYIFLFTVLLSEFTLTAPPCCGMTSSSFYQEFLLRMQRPQIDAPSTRSLSKGP 60
OY 61 TETFAHMPRNCYHSATLTCMHANTHWTKMKNPSCPGIGVYVCFTFTQGMMSGGV 120
    |||||||
DB 61 TETFAHMPRNCYHSATLTCMHANTHWTKMKNPSCPGIGVYVCFTFTQGMMSGGV 120
OY 121 ODQAREKHVKEVISQLTRVHGTSPPYKGLDSLKLTHTLTHRLVSLFTTTLGLHEVSA 180
    |||||||
DB 121 ODQAREKHVKEVISQLTRVHGTSPPYKGLDSLKLTHTLTHRLVSLFTTTLGLHEVSA 180
OY 181 QNEPTNCWICLPINFREYVSIYPVEQWNNFSTEINTTSVLVGLVLSNLETHRSNLTGVKF 240
    |||||||
DB 181 QNEPTNCWICLPINFREYVSIYPVEQWNNFSTEINTTSVLVGLVLSNLETHRSNLTGVKF 240
OY 241 SNTTYTTSNQCIRWVPPPOIYCLPSGIFVCGTSAVRLNGSSSMCFSLFVPMPTIY 300
    |||||||
DB 241 SNTTYTTSNQCIRWVPPPOIYCLPSGIFVCGTSAVRLNGSSSMCFSLFVPMPTIY 300
OY 301 TEODLXSYISKPRNKRVPILPVIAGVLAGLTGIGITTSSTOFYKLSOELANDMER 360
    |||||||
DB 301 TEODLXSYISKPRNKRVPILPVIAGVLAGLTGIGITTSSTOFYKLSOELANDMER 360
OY 361 VADSLVTLDQDLSLAAYVQLQNRALDILLTAERGCTCLFGECCYYVQSGIVTEKYKE 420
    |||||||
DB 361 VADSLVTLDQDLSLAAYVQLQNRALDILLTAERGCTCLFGECCYYVQSGIVTEKYKE 420
OY 421 IRDRIQRAAEELRNTGPMGLLSQMPMPILPFLGPIAAIILLFGPCIFNLVNFVSSRI 480
    |||||||
DB 421 IRDRIQRAAEELRNTGPMGLLSQMPMPILPFLGPIAAIILLFGPCIFNLVNFVSSRI 480
OY 481 EAVKLOMEPRKMSKTKIYRRPLDRPASPSDVNDIKGTPPEEISAQPLLRPSAGSS 538
    |||||||
DB 481 EAVKLOMEPRKMSKTKIYRRPLDRPASPSDVNDIKGTPPEEISAQPLLRPSAGSS 538

RESULT 7
AA67313
ID AAY67313 standard; Protein; 538 AA.
AC AAY67313;
XX
XX 11-APR-2000 (first entry)
XX
XX Human secreted protein AJ172_2 amino acid sequence.
XX
XX Homo sapiens.
XX
XX Human; secreted protein; disease diagnosis; pre-eclampsia; cancer;
KW placental pathology; metastasis inhibition; nutritional activity;
KW immune stimulator; haematopoiesis regulator; tissue growth;
KW tumour inhibitor; anti-inflammatory; clone AJ172_2; ATCC_98115.
XX
XX OS
XX
XX WO9960020-A1.
XX
XX 25-NOV-1999.
XX
XX 17-MAY-1999; 99NO-US10915.
XX
XX 18-MAY-1998; 98US-0080478.
XX
XX 20-OCT-1998; 98US-0175928.
XX
XX (GENE ) GENETICS INST INC.
XX
XX Jacobs K, McCoy JM, Lavallic ER, Collins-Racie LA, Evans C;
XX
XX Werberg D, M S, Treacy M;
XX
XX WPI; 2000-116311/10.

```

DR N-PSDB: AA259468.

XX New polynucleotides encoding secreted cDNA libraries, used to develop
PT products for the diagnosis and treatment of neoplastic disease

XX Claim 15: Page 108-110; 149pp; English.

XX This is the human secreted protein A172.2 amino acid sequence, the
CC polynucleotide encoding A172.2 was obtained from a human adult testes
CC cDNA library. The invention relates to secreted human and murine
CC proteins. The polynucleotides and proteins are predicted to have
CC biological activities which would make them suitable for treating,
CC preventing or ameliorating medical conditions in humans and animals.
CC Detection of the levels of the proteins can be used for the diagnosis of
CC e.g. pre-eclampsia, placental pathology or cancer. Agents which modulate
CC the expression or function of the proteins may be used for treating a
CC neoplastic disease and inhibiting metastasis. Other suggested activities
CC include nutritional activity (e.g. in feeds), cytokine and cell
CC proliferation/differentiation activity, immune stimulating (e.g. as
CC vaccines) or suppressing activity, haematopoiesis regulating activity,
CC tissue growth activity, activin/inhibin activity,
CC chemotactic/chemokinetic activity, haemostatic and thrombolytic activity,
CC receptor/ligand activity, anti-inflammatory activity, cadherin/tumour
CC invasion suppressor activity, and tumour inhibition activity. The
CC polynucleotide sequences are also stated to be useful for gene therapy.

XX Sequence 538 AA:

Query Match 99.9%; Score 2868; DB 21; Length 538;
Best Local Similarity 99.8%; Pred. No. 5.9e-259;
Matches 537; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALPYHIFLFTVLLPSFTLTAPPCRCMTSSPYOEFLMRQRCNDAPSYRSLSKCTP 60
DB 1 MALPYHIFLFTVLLPSFTLTAPPCRCMTSSPYOEFLMRQRCNDAPSYRSLSKCTP 60
QY 61 TFTAHTHPRNCYSATLCMHANTHYWTKMINSCPGGLGVTCWTFYTGMSDGGV 120
DB 61 TFTAHTHPRNCYSATLCMHANTHYWTKMINSCPGGLGVTCWTFYTGMSDGGV 120
QY 121 QDQAREKHVKVVISQLTRVHGTSSPYKGLDLSKHETLRTHTRLVSLFNTTLTGHEVSA 180
DB 121 QDQAREKHVKVVISQLTRVHGTSSPYKGLDLSKHETLRTHTRLVSLFNTTLTGHEVSA 180
QY 181 QNPNCWICLP LNRPVYSIPVPEOMNFEINNTSVLGPVLSNLEITHTSNLTCKVF 240
DB 181 QNPNCWICLP LNRPVYSIPVPEOMNFEINNTSVLGPVLSNLEITHTSNLTCKVF 240
QY 241 SNTTYTNSOCIRWVTPPTQIVCLPSGIFVCGTSAYRCLNGSSSESMCFSLFVPPMTIY 300
DB 241 SNTTYTNSOCIRWVTPPTQIVCLPSGIFVCGTSAYRCLNGSSSESMCFSLFVPPMTIY 300
QY 301 TEODLYSVISKPRNKRVPILPFVIGAGVLCALGTGIGITTSQFYKLSQELNGDME 360
DB 301 TEODLYSVISKPRNKRVPILPFVIGAGVLCALGTGIGITTSQFYKLSQELNGDME 360
QY 361 VADSLVTLQDGLNLAVALVONRRALDLTLAERGTCFLGEECCYVYVNOGITYEYKE 420
DB 361 VADSLVTLQDGLNLAVALVONRRALDLTLAERGTCFLGEECCYVYVNOGITYEYKE 420
QY 421 IRDRIORRAEELRNTGPMGLSOMMPWILPFLGPLAAIILLLLEFCPIFNLLVYVSSRI 480
DB 421 IRDRIORRAEELRNTGPMGLSOMMPWILPFLGPLAAIILLLLEFCPIFNLLVYVSSRI 480
QY 481 EAVKLOWEPKMQSTKTIYRRLDPPASPRSDVNDIKGTPPEEISAAPLLRPNSAGSS 538
DB 481 EAVKLOWEPKMQSTKTIYRRLDPPASPRSDVNDIKGTPPEEISAAPLLRPNSAGSS 538

RESULT 8
AAE14540
ID AAE14540 standard; Protein; 538 AA.

AC AAE14540:

XX 07-MAY-2002 (first entry)

XX Human syncytin protein.

XX Human; syncytin; pre-eclampsia; gestational trophoblast disorder;
KW choriocarcinoma; hydatiform mole; placental site tumour; abortion;
KW envelope gene; human endogenous defective retrovirus; HERV-W.

XX Homo sapiens.

XX WO200204678-A2.

XX 17-JAN-2002.

XX 09-JUL-2001; 2001WO-US21719.

XX 07-JUL-2000; 2000US-216657P.

XX (GENY) GENETICS INST INC.

XX Keith JC, McCoy JM, MI S;

XX WPI: 2002-171727/22.

XX N-PSDB: AAD24195.

PT Identifying a compound for treating a subject with or at risk of
PT developing pre-eclampsia, comprises determining whether the expression
PT or activity of syncytin in the cell is modulated in the presence of a
PT test compound

PS Disclosure: Page 42-43; 43pp; English.

XX The invention relates to identifying compounds which are modulators
CC of syncytin expression. The syncytin modulators are useful in diagnosis
CC and treatment of pre-eclampsia and gestational trophoblast disorders (e.g.
CC choriocarcinoma, hydatiform mole, placental site tumour and missed/
CC incomplete abortion). Syncytin is a human gene derived from the
CC envelope gene of human endogenous defective retrovirus, HERV-W. The
CC present invention is based partly on the discovery that syncytin
CC expression is dramatically reduced in pre-eclampsia, and is also
CC mis-localised to the apical syncytiotrophoblast membrane. The present
CC sequence is human syncytin protein.

XX Sequence 538 AA:

Query Match 99.9%; Score 2868; DB 23; Length 538;
Best Local Similarity 99.8%; Pred. No. 5.9e-259;
Matches 537; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALPYHIFLFTVLLPSFTLTAPPCRCMTSSPYOEFLMRQRCNDAPSYRSLSKCTP 60
DB 1 MALPYHIFLFTVLLPSFTLTAPPCRCMTSSPYOEFLMRQRCNDAPSYRSLSKCTP 60
QY 61 TFTAHTHPRNCYSATLCMHANTHYWTKMINSCPGGLGVTCWTFYTGMSDGGV 120
DB 61 TFTAHTHPRNCYSATLCMHANTHYWTKMINSCPGGLGVTCWTFYTGMSDGGV 120
QY 121 QDQAREKHVKVVISQLTRVHGTSSPYKGLDLSKHETLRTHTRLVSLFNTTLTGHEVSA 180
DB 121 QDQAREKHVKVVISQLTRVHGTSSPYKGLDLSKHETLRTHTRLVSLFNTTLTGHEVSA 180
QY 181 QNPNCWICLP LNRPVYSIPVPEOMNFEINNTSVLGPVLSNLEITHTSNLTCKVF 240
DB 181 QNPNCWICLP LNRPVYSIPVPEOMNFEINNTSVLGPVLSNLEITHTSNLTCKVF 240
QY 241 SNTTYTNSOCIRWVTPPTQIVCLPSGIFVCGTSAYRCLNGSSSESMCFSLFVPPMTIY 300
DB 241 SNTTYTNSOCIRWVTPPTQIVCLPSGIFVCGTSAYRCLNGSSSESMCFSLFVPPMTIY 300
QY 301 TEODLYSVISKPRNKRVPILPFVIGAGVLCALGTGIGITTSQFYKLSQELNGDME 360
DB 301 TEODLYSVISKPRNKRVPILPFVIGAGVLCALGTGIGITTSQFYKLSQELNGDME 360

Db 301 TEODLYNYISKRPNKRVPIILPFVIGAGVIGALGTGIGITTTSTOFYKXLSOELNDMER 360
 QY 361 VADSLVTTLODQNSLAAYVLONRRAIDLTAERGCTCLFLGEECCYVNOGSIYVEKVE 420
 Db 361 VADSLVTTLODQNSLAAYVLONRRAIDLTAERGCTCLFLGEECCYVNOGSIYVEKVE 420
 QY 421 IDRIGORAEELRNTGPMGLLSQMPMILPFLGPIAAITLLLFGPCIFNLLVNFVSSRI 480
 Db 421 IDRIGORAEELRNTGPMGLLSQMPMILPFLGPIAAITLLLFGPCIFNLLVNFVSSRI 480
 QY 481 EAVKLQMEPKMOSKTKIYRRPLDRPASPRSDVNDIKGTPPEEISAQPLLRNSAGSS 538
 Db 481 EAVKLQMEPKMOSKTKIYRRPLDRPASPRSDVNDIKGTPPEEISAQPLLRNSAGSS 538

RESULT 9
 ID AAM99345 standard; Protein: 538 AA.
 XX AAM99345;

XX 21-MAY-1999 (first entry)

XX Human endogenous retrovirus W ORF1.

XX Clone: human endogenous retrovirus; genome; autoimmune disease;
 KM multiple sclerosis; rheumatoid polyarthritis; insulin-dependent diabetes;
 KM disseminated lupus erythematosus; pregnancy; chromosomal marker.

XX Human endogenous retrovirus.

XX WO9902696-A1.

XX 21-JAN-1999.

XX 06-JUL-1998; 98WO-FR01442.

XX 07-JUL-1997; 97FR-0008815.

XX (INMR) BIO MERIEUX.

XX Beseme F, Blond JL, Bouton O, Mallet F, Mandrand B.

XX WPI; 1999-120897/10.

XX New nucleic acid sequences from human endogenous retrovirus-W -
 PT expressed exclusively in placenta and useful in diagnosis and
 PT therapy of autoimmune disease, and abnormal or failed pregnancy

XX Example 8; Fig 6; 106bp; French.

XX This sequence represents the open reading frame 1 encoding the env
 CC protein from the human endogenous retrovirus (HENV) W (AA25665).
 CC Nucleic acids, their fragments or peptides encoded by them derived from
 CC the HENV-W genome are markers of autoimmune disease (e.g. multiple
 CC sclerosis, rheumatoid polyarthritis, disseminated lupus erythematosus,
 CC insulin-dependent diabetes and related pathologies) and of abnormal or
 CC unsuccessful pregnancy and can be used as chromosomal markers for
 CC susceptibility to these conditions, or proximity markers of genes
 CC associated with this susceptibility.

XX SQ Sequence 538 AA;

Query Match 97.2%; Score 2791; DB 20; Length 538;
 Best Local Similarity 97.2%; Pred. No. 9, 4e-252;
 Matches 523; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 MALPYHIFLFTVLPSEFTLAPPCRCMTSSPYQEFILMRQPGNIDAPSYRSLSKGP 60
 Db 1 MGLPYHIFLCSVLSPCFTLAPPCRCMTSSPHPEFLMRQPGNIDAPSYRSLSKGP 60
 QY 61 TFTAHTHMRNCGYHSATLCMHANTHYTGKMINPSCPGGLGYVVCWTFYTGQMSDGGCV 120

Db 61 TFTAHTHMRNCGYHSATLCMHANTHYTGKMINPSCPGGLGYVVCWTFYTGQMSDGGCV 120
 QY 121 QDAREKHVEKYSQLTRVHGTSPPYKGLDLSKLTHTLRRLVSLFNTTLGLHEVSA 180
 Db 121 QDAREKHVEKYSQLTRVHGTSPPYKGLDLSKLTHTLRRLVSLFNTTLGLHEVSA 180
 QY 181 QNPNTNCWICLPNFRPYVSIIPVEQWNNFSTEINTTSVYVGPLVSNLETHTNSNLTCVFR 240
 Db 181 QNPNTNCWICLPNFRPYVSIIPVEQWNNFSTEINTTSVYVGPLVSNLETHTNSNLTCVFR 240
 QY 241 SMTYTTNSQCIRWVTPPTQIVCLPSGIFVCGTSAVRCNLNGSSSEMCFLSFLVPMYTY 300
 Db 241 SMTYTTNSQCIRWVTPPTQIVCLPSGIFVCGTSAVRCNLNGSSSEMCFLSFLVPMYTY 300
 QY 301 TEODLYSYISKRPNKRVPIILPFVIGAGVIGALGTGIGITTTSTOFYKXLSOELNDMER 360
 Db 301 TEODLYSYISKRPNKRVPIILPFVIGAGVIGALGTGIGITTTSTOFYKXLSOELNDMER 360
 QY 361 VADSLVTTLODQNSLAAYVLONRRAIDLTAERGCTCLFLGEECCYVNOGSIYVEKVE 420
 Db 361 VADSLVTTLODQNSLAAYVLONRRAIDLTAERGCTCLFLGEECCYVNOGSIYVEKVE 420
 QY 421 IDRIGORAEELRNTGPMGLLSQMPMILPFLGPIAAITLLLFGPCIFNLLVNFVSSRI 480
 Db 421 IDRIGORAEELRNTGPMGLLSQMPMILPFLGPIAAITLLLFGPCIFNLLVNFVSSRI 480
 QY 481 EAVKLQMEPKMOSKTKIYRRPLDRPASPRSDVNDIKGTPPEEISAQPLLRNSAGSS 538
 Db 481 EAVKLQMEPKMOSKTKIYRRPLDRPASPRSDVNDIKGTPPEEISAQPLLRNSAGSS 538

RESULT 10
 ID AAE25054 standard; Protein: 542 AA.
 XX AAE25054;

XX 30-OCT-2002 (first entry)

XX Human EMBRY-2 protein.

XX Human; embryogenesis associated protein; AIDS; reproductive disorder;
 KM infertility; endometriosis; endometrial tumour; inflammatory disorder;
 KM autoimmune disorder; acquired immune deficiency syndrome; transgenic;
 KM ovarian tumour; contact dermatitis; placenta disorder; preclampsia;
 KM EMBRY-2; allergy; gene therapy.

XX Homo sapiens.

XX Key Location/Qualifiers

XX Peptide 1..20

XX Protein 21..542 /label= Signal-peptide

XX Domain 263..484 /note= "Mature EMBRY-2 protein"

XX Domain 455..473 /note= "ENV polypeptide (coat polypeptide) domain"

XX WO200248362-A2.

XX 20-JUN-2002.

XX 14-NOV-2001; 2001WO-US43956.

XX 15-NOV-2000; 2000US-249407P.

XX (INCY-) INCYTE GENOMICS INC.

XX Rankumar J, Arvizu C.

XX WPI; 2002-537629/57.

XX N-PSDB; AAD41225.

XX New polypeptides of human embryogenesis associated proteins for
 PT screening modulators useful for treating or preventing disorders e.g.
 PT endometriosis, infertility, allergy, preclampsia

PS Claim 57; Page 94-95; 97pp; English.

XX The invention relates to human embryogenesis associated proteins
 CC (EMBR) and nucleic acid molecules encoding such proteins. EMBRY
 CC sequences are useful for screening modulators useful for treating
 CC or preventing disorders associated with abnormal expression of
 CC EMBRY. The disorders treated include reproductive disorders such
 CC as infertility, endometriosis, endometrial or ovarian tumour;
 CC autoimmune/inflammatory disorder such as acquired immune deficiency
 CC syndrome (AIDS), allergies, contact dermatitis; disorders of the
 CC placenta such as preclampsia, abruptio placentae etc. Sequences
 CC of the invention are also useful for analysing a proteome of a
 CC tissue or a cell type. EMBRY proteins are useful as immunogens for
 CC preparing antibodies. Polynucleotides of the invention are useful
 CC for creating knockin humanised animals or transgenic animals to
 CC model human diseases. They are also used in gene therapy. The
 CC present sequence is human EMBRY-2 protein.

XX Sequence 542 AA:

Query Match 89.1%; Score 2558; DB 23; Length 542;

Best Local Similarity 88.4%; Pred. No. 5,86-230;

Matches 479; Conservative 22; Mismatches 37; Indels 4; Gaps 1;

QY 1 MALPYHIFLFTVLLPSTLTAPPCRCMTSSSPYQEFILMRQPGNIDAPSRSLSGTP 60
 DB 1 MALPYCIFLFTVLLPSTLTAPPCRCMTSSSPYQAEILMRMRPHIDAPSRSLSGNP 60
 QY 1 TTTAHTMPRCYASATLCMHANTHYMTGKMINSPCGGLGVTVCMTYFTQTGSDGCV 120
 DB 61 AFTAHTMPNCISATLCMHANTHYMTGKMINSPCGGLGATTCMTYFTQTGSDGCV 120
 QY 121 QDQAREKHVEVISQLTRVHGTSSPYKGLDLSKHELTHTRTVSLFNTLTGLHEVSA 180
 DB 121 QDQAREKHVEVISQLTRVHGTSSPYKGLDLSKHELTHTRTVSLFNTLTGLHEVSA 180
 QY 181 QNPNCNICLPNRPVSLIPVPCOMNNESTEITTSVGLVPLVSNLEITHTSNLTCVKF 240
 DB 181 QNPNCNICLPNRPVSLIPVPCOMNNESTEITTSVGLVPLVSNLEITHTSNLTCVKF 240
 QY 241 SNTTNTSOCIIRVPTPTOIVCLPSGIFPCVGTSAVRCNLGSSSCFSLVPPMTIY 300
 DB 241 SNTTNTSOCIIRVPTPTOIVCLPSGIFPCVGTSAVRCNLGSSSCFSLVPPMTIY 300
 QY 301 TEODLSYVISKPRNKRVPILPFVIGAGVLGALGTGIGITSTQFYKLSQELNCDMER 360
 DB 301 TEODLSYVISKPRNKRVPILPFVIRAGVIGRLGTGISTSTQFYKLSQELNCDMER 360
 QY 361 VADSUVLTQDQNLAVLQNRALDLTAERGGTCLFGEDECCYVNOGIVTEVYKE 420
 DB 361 VTSLSVLTQDQNLAVLQNRALDLTAERGGTCLFGEDECCYVNOGIVTEVYKE 420
 QY 421 IRDRIORAEELRTGPMGLLSOMPWITLPLGLALAILLLGPCIENLTVFVSRI 480
 DB 421 IRDRIORAEELRTGPMGLLSOMPWITLPLGLALAILLLGPCIENLTVFVSRI 480
 QY 481 EAVK---LQMEPKMQSKTIYRRPLDRPASPRSDVNDIKGTPEEISAAQPLLRPSAG 536
 DB 481 EAVK---LQMEPKMQSKTIYRRPLDRPASPRSDVNDIKGTPEEISAAQPLLRPSAG 536
 QY 537 SS 538
 DB 541 SS 542

RESULT 11
 AAB08195
 ID AAB08195 standard; Protein; 541 AA.

XX AAB08195;
 AC 04-DEC-2000 (first entry)
 XX
 DT
 DE Amino acid sequence of the MSRV-1 3' env and LTR regions.
 XX
 KM MSRV-1; pol region; long terminal repeat; LTR; RUS region; retrovirus.
 XX
 OS Multiple Sclerosis retrovirus 1.

XX Key
 FH Peptide
 FT location/Qualifiers
 FT 1..27
 FT /note= "signal peptide"

FT Modified-site
 FT /note= "N-linked glycosylation site"

FT Modified-site
 FT /note= "N-linked glycosylation site"

FT Modified-site
 FT /note= "N-linked glycosylation site"

FT Modified-site
 FT /note= "N-linked glycosylation site"

FT Modified-site
 FT /note= "N-linked glycosylation site"

FT Modified-site
 FT /note= "N-linked glycosylation site"

FT Cleavage-site
 FT /note= "putative cleavage site"

FT Cleavage-site
 FT /note= "putative cleavage site"

FT Modified-site
 FT /note= "N-linked glycosylation site"

FT Modified-site
 FT /note= "N-linked glycosylation site"

FT Modified-site
 FT /note= "N-linked glycosylation site"

FT Modified-site
 FT /note= "N-linked glycosylation site"

FT Modified-site
 FT /note= "N-linked glycosylation site"

FT Modified-site
 FT /note= "N-linked glycosylation site"

FT Modified-site
 FT /note= "N-linked glycosylation site"

FT Modified-site
 FT /note= "N-linked glycosylation site"

FT Modified-site
 FT /note= "N-linked glycosylation site"

FT Modified-site
 FT /note= "N-linked glycosylation site"

FT Modified-site
 FT /note= "N-linked glycosylation site"

FT Modified-site
 FT /note= "N-linked glycosylation site"

FT Modified-site
 FT /note= "N-linked glycosylation site"

FT Modified-site
 FT /note= "N-linked glycosylation site"

FT Modified-site
 FT /note= "N-linked glycosylation site"

FT Modified-site
 FT /note= "N-linked glycosylation site"

FT Modified-site
 FT /note= "N-linked glycosylation site"

FT Modified-site
 FT /note= "N-linked glycosylation site"

FT Modified-site
 FT /note= "N-linked glycosylation site"

FT Modified-site
 FT /note= "N-linked glycosylation site"

FT Modified-site
 FT /note= "N-linked glycosylation site"

FT Modified-site
 FT /note= "N-linked glycosylation site"

FT Modified-site
 FT /note= "N-linked glycosylation site"

FT Modified-site
 FT /note= "N-linked glycosylation site"

FT Modified-site
 FT /note= "N-linked glycosylation site"

FT Modified-site
 FT /note= "N-linked glycosylation site"

```

Db      60 TTTAHTHMRNCYNSATLCMHANTHYWTKMINPSCPGGLGATVCWTYFTHTSMSDGGI 119
QY      121 QDOAREKHVKEVISQLTRHGHSSPYKGLDLSKHHTLTHTRLVSLFNTTTLTGLHEVA 180
Db      120 QGOAREKOVKEAISQLTRGHSTSPYKGLVLSKLTHTLTHRLVSLFNTTTLRLHEVA 179
QY      181 QNPNCWICLPINFRPYVSIPEEOMNNESTEINTTSVYGLVSNLETHHSNLTGVCF 240
Db      180 QNPNCWICLPINFRPYVSIPEEOMNNESTEINTTSVYGLVSNLETHHSNLTGVCF 239
QY      241 SNTTYTNSQCIKRWTPPTQIVCLPSGIFVCGTSAYRCLNGSSSFCFLSLVPMPTIY 300
Db      240 SNTIDTSSQCIKRWTPPTQIVCLPSGIFVCGTSAYRCLNGSSSFCFLSLVPMPTIY 299
QY      301 TEODLSYVYSKPRKRVPLPFPVIGAGVLGAGTGIGITSTQFYKLSOELNDMER 360
Db      300 TEODLXNHVYKPRKRVPLPFPVIRAGVLRGIGITSTQFYKLSOELNDMEQ 359
QY      361 VADSLVTLQDQNSLAAYVLAQNRALDLTLTAERGTCFLGEECCYVNSGIVTEKYK 420
Db      360 VTDSLVTTLQDQNSLAAYVLAQNRALDLTLTAERGTCFLGEECCYVNSGIVTEKYK 419
QY      421 IRDRIOGRAEELRNTGPMGLSQMPWILPGLPLAIIILLFGPCIFNLLVNFVSSRI 480
Db      420 IRDRIOGRAEELQNTERMGILSQMPWVLPFLGPLAAIILLFGPCIFNLLVNFVSSRI 479
QY      481 EAVK----LQMEPKMOSKTKIYRRPLDRPASPSRDVNDIKGTPPEISAAOPLLRNSAG 536
Db      480 EAVKIQVLMQMEPOMESMTKIHGRPLDRPASPCSDVNDIEGTPPEISTAOPLLRNSAG 539
QY      537 SS 538
Db      540 SS 541

RESULT 12
AAB19069
ID      AAB19069 standard; Protein: 542 AA.
XX
AC      AAB19069;
XX
DT      08-FEB-2001 (first entry)
XX
DE      Amino acid sequence of an envelope (env) protein of MSRV-1.
XX
KW      Envelope protein; MSRV-1; superantigen; autoimmune disease; Vbeta16;
KM      Vbeta17; multiple sclerosis; vaccine.
XX
OS      Human endogenous retrovirus.
XX
PN      WO200057185-A1.
XX
PD      28-SEP-2000.
XX
PF      20-MAR-2000; 2000WO-FR00691.
XX
PR      19-MAR-1999; 99FR-0003622.
XX
PR      28-OCT-1999; 99FR-0013755.
XX
PA      (INMR ) BIO MERIEUX.
XX
PI      Perron H, Lafont M;
XX
DR      MPI: 2000-638379/61.
XX
DR      N-PSDB: AAA96625.
XX
PT      Detecting superantigen activity, useful for identifying agents for
XX      treatment or prevention of autoimmune disease, from expansion or loss
XX      of particular lymphocyte Vbeta determinants
PS      Claim 28; Page 124-126; 134pp; French.
XX
CC      The present sequence represents an envelope protein of MSRV-1. The

```

```

CC      envelope protein expressed by the endogenous human retrovirus MSRV-1
CC      has superantigen activity associated with autoimmune disease. The
CC      protein can be detected using the method of the invention. The
CC      specification describes a process for detecting activity of a
CC      superantigen in a biological sample. The process comprises identifying
CC      a large scale expansion or loss of lymphocytes that carry at least one
CC      of the Vbeta16 and/or Vbeta17 determinants. The method is used to
CC      screen for agents that inhibit the superantigen, especially those
CC      associated with MSRV-1 which is implicated in autoimmune disease.
CC      particularly multiple sclerosis. These agents are potentially useful
CC      for treatment or prevention (e.g. as vaccines) of autoimmune diseases.
XX
SQ      Sequence      542 AA:
Query Match      86.7%; Score 2490; DB 21; Length 542;
Best Local Similarity 87.3%; Pred. No. 1.3e-223;
Matches 473; Conservative 17; Mismatches 48; Indels 4; Gaps 1;

QY      1 MALPYHIFLFTVLLSFYTLAPPCRCMTSSPYQEFRLRMORPGNIDAPYSRLSKGTP 60
Db      1 MALPYHIFLFTVLLPFPALTAAPPCCTTSSSPYQEFRLMRPLPGNIDAPYSRLSKGNS 60
QY      61 TTTAHTHMRNCYNSATLCMHANTHYWTKMINPSCPGGLGATVCWTYFTHTSMSDGGI 120
Db      61 TTTAHTHMRNCYNSATLCMHANTHYWTKMINPSCPGGLGATVCWTYFTHTSMSDGGI 120
QY      121 QDOAREKHVKEVISQLTRHGHSSPYKGLDLSKHHTLTHTRLVSLFNTTTLTGLHEVA 180
Db      121 QGOAREKOVKEAISQLTRGHSTSPYKGLVLSKLTHTLTHRLVSLFNTTTLRLHEVA 180
QY      181 QNPNCWICLPINFRPYVSIPEEOMNNESTEINTTSVYGLVSNLETHHSNLTGVCF 240
Db      181 QNPNCWICLPINFRPYVSIPEEOMNNESTEINTTSVYGLVSNLETHHSNLTGVCF 240
QY      241 SNTTYTNSQCIKRWTPPTQIVCLPSGIFVCGTSAYRCLNGSSSFCFLSLVPMPTIY 300
Db      241 SNTIDTSSQCIKRWTPPTQIVCLPSGIFVCGTSAYRCLNGSSSFCFLSLVPMPTIY 300
QY      301 TEODLSYVYSKPRKRVPLPFPVIGAGVLGAGTGIGITSTQFYKLSOELNDMER 360
Db      301 TEODLXNHVYKPRKRVPLPFPVIRAGVLRGIGITSTQFYKLSOELNDMEQ 360
QY      361 VADSLVTLQDQNSLAAYVLAQNRALDLTLTAERGTCFLGEECCYVNSGIVTEKYK 420
Db      361 VTDSLVTTLQDQNSLAAYVLAQNRALDLTLTAERGTCFLGEECCYVNSGIVTEKYK 420
QY      421 IRDRIOGRAEELRNTGPMGLSQMPWILPGLPLAIIILLFGPCIFNLLVNFVSSRI 480
Db      421 IRDRIOGRAEELQNTERMGILSQMPWVLPFLGPLAAIILLFGPCIFNLLVNFVSSRI 480
QY      481 EAVK----LQMEPKMOSKTKIYRRPLDRPASPSRDVNDIKGTPPEISAAOPLLRNSAG 536
Db      481 EAVKIQVLMQMEPOMESMTKIHGRPLDRPARLCSVDVNDIEGTPPEISTAOPLLRNSVAG 540
QY      537 SS 538
Db      541 SS 542

RESULT 13
AAW71068
ID      AAW71068 standard; Protein: 493 AA.
XX
AC      AAW71068;
XX
DT      29-DEC-1998 (first entry)
XX
DE      Multiple sclerosis associated retrovirus protein 3.
XX
KW      Multiple sclerosis associated retrovirus; MSRV; MS; pol gene;
KM      gag gene; env gene; rheumatoid arthritis-associated virus.
XX
OS      Multiple sclerosis associated retrovirus.

```

```

XX Key Location/Qualifiers
FH Misc-difference 39 /note- "Encoded by tga"
FT XX
XX
XX W09823755-A1.
XX
XX 04-JUN-1998.
XX
XX 26-NOV-1997; 97MO-IB01482.
XX
XX 26-NOV-1996; 96US-0756429.
XX
XX (INMR ) BIO MERIEUX.
XX
XX Bedin F, Beseme F, Jolivet-Reynaud C, Komurian-Pradel F;
PI Mandrand B, Paranhos-Baccala G, Perron H;
DR WPI; 1998-322732/28.
DR N-PSDB; AAV43217.
XX
XX New nucleic acid from retroviruses - useful for diagnosis,
PT prevention and treatment of, e.g. multiple sclerosis
XX
XX Disclosure; Pages 185-187; 286pp; English.
XX
XX The present sequence represents a multiple sclerosis (MS) associated
CC retrovirus (MSRV) protein used in the method of the invention. The
CC invention provides complete or partial genomic sequences of the
CC MSRV-1 pol gene, gag gene and env gene, and polypeptides encoded by
CC these genes. The invention also provides antibodies raised against
CC the polypeptides. The genomic sequences, polypeptides and antibodies
CC are also claimed useful for diagnosing infection by MS and rheumatoid
CC arthritis-associated viruses, and also for prevention and treatment of
CC infection with these viruses.
XX
XX Sequence 493 AA;
SQ
Query Match 80.6%; Score 2314; DB 19; Length 493;
Best Local Similarity 88.4%; Pred. No. 3.3e-207;
Matches 435; Conservative 19; Mismatches 38; Indels 0; Gaps 0;

QY 1 MALPYHIFLFTVLLPSTLTAPPCRCMTSSSPYQEFLLMRQRPNDASYSRLSKGTP 60
DB 1 MALPYHIFLFTVLLPSTLTAPPCRCMTSSSPYQEFLLMRQRPNDASYSRLSKGNS 60
QY 61 TFTAHTMPRNCYHSATLCMHANTHYWTGKMINSPSCGGLGVYVCWTFYQTGMSDGGV 120
DB 61 TFTAHTMPRNCYHSATLCMHANTHYWTGKMINSPSCGGLGVYVCWTFYQTGMSDGGI 120
QY 121 ODQAREKHVEVISQTLRVHGTSPPYKGLDSLKHETLRHTRLVSLFNTTLTGLHEVSA 180
DB 121 ODQAREKHVEVISQTLRVHGTSPPYKGLDSLKHETLRHTRLVSLFNTTLTGLHEVSA 180
QY 181 QNPNCMICPLNFRPVYSIPVPEQNNFSTEINTTSVLVGPLVSNLEITHSMLTCVKF 240
DB 181 QNPNCMICPLNFRPVYSIPVPEQNNFSTEINTTSVLVGPLVSNLEITHSMLTCVKF 240
QY 241 SNTTYTNSOCIRWVPTQIVCLPSGIFVCGSAVRCUNGSSSCFSLVPPMTIY 300
DB 241 SNTTYTNSOCIRWVPTQIVCLPSGIFVCGSAVRCUNGSSSCFSLVPPMTIY 300
QY 301 TEODLVYSVSKPRNKVPLIPVIGAGVLGALGTGIGITTSQTFYKLSQELGDMER 360
DB 301 TEODLVYSVSKPRNKVPLIPVIGAGVLGALGTGIGITTSQTFYKLSQELGDMER 360
QY 361 VADSLVTLQDQLSLAAVVLQNRALDGLTAERGGTCLTGECCYYVNSGIVTEKVK 420
DB 361 VADSLVTLQDQLSLAAVVLQNRALDGLTAERGGTCLTGECCYYVNSGIVTEKVK 420
QY 421 IRRIORRAEELNNTGWSGLSOWMPWILPFLGPIAAIILLFGPCEIFMLVNFVSSRI 480
DB 421 IRRIORRAEELNNTGWSGLSOWMPWILPFLGPIAAIILLFGPCEIFMLVNFVSSRI 480

```

```

QY 481 EAVKLOMEPKMO 492
DB 481 EAVKLOMYLOME 492

RESULT 14
ID AAM9553
XX AAM9553 standard; Protein: 493 AA.
XX
XX AAM9553;
XX
XX 08-JUN-1999 (first entry)
XX
XX Protein encoded by clone C15 from MSRV-1.
XX
XX Multiple sclerosis; virus; diagnosis; prophylaxis; therapy; expression;
XX rheumatoid polyarthritis.
XX
XX Multiple sclerosis related virus type 1.
XX
XX Key Location/Qualifiers
FH Misc-difference 39 /label- unknown
FT XX
XX FR2765588-A1.
XX
XX 08-JAN-1999.
XX
XX 07-JUL-1997; 97FR-0008816.
XX
XX 07-JUL-1997; 97FR-0008816.
XX
XX (INMR ) BIO MERIEUX.
XX
XX WPI; 1999-098275/09.
XX
XX N-PSDB; AAX29703.
XX
XX Nucleic acid sequences of retrovirus called MSRV-1 - associated with
PT multiple sclerosis or rheumatoid polyarthritis
XX
XX Claim 2; Page 38-39; 83pp; French.
XX
XX This sequence represents the protein encoded by clone C15 from a
CC novel multiple sclerosis related virus type 1 (MSRV1). The sequence
CC can be used in diagnostic, prophylactic or therapeutic compositions
CC to inhibit expression of a multiple sclerosis related virus and/or
CC virus associated with rheumatoid polyarthritis.
XX
XX Sequence 493 AA;
SQ
Query Match 80.6%; Score 2314; DB 20; Length 493;
Best Local Similarity 88.4%; Pred. No. 3.3e-207;
Matches 435; Conservative 19; Mismatches 38; Indels 0; Gaps 0;

QY 1 MALPYHIFLFTVLLPSTLTAPPCRCMTSSSPYQEFLLMRQRPNDASYSRLSKGTP 60
DB 1 MALPYHIFLFTVLLPSTLTAPPCRCMTSSSPYQEFLLMRQRPNDASYSRLSKGNS 60
QY 61 TFTAHTMPRNCYHSATLCMHANTHYWTGKMINSPSCGGLGVYVCWTFYQTGMSDGGV 120
DB 61 TFTAHTMPRNCYHSATLCMHANTHYWTGKMINSPSCGGLGVYVCWTFYQTGMSDGGI 120
QY 121 ODQAREKHVEVISQTLRVHGTSPPYKGLDSLKHETLRHTRLVSLFNTTLTGLHEVSA 180
DB 121 ODQAREKHVEVISQTLRVHGTSPPYKGLDSLKHETLRHTRLVSLFNTTLTGLHEVSA 180
QY 181 QNPNCMICPLNFRPVYSIPVPEQNNFSTEINTTSVLVGPLVSNLEITHSMLTCVKF 240
DB 181 QNPNCMICPLNFRPVYSIPVPEQNNFSTEINTTSVLVGPLVSNLEITHSMLTCVKF 240
QY 241 SNTTYTNSOCIRWVPTQIVCLPSGIFVCGSAVRCUNGSSSCFSLVPPMTIY 300
DB 241 SNTTYTNSOCIRWVPTQIVCLPSGIFVCGSAVRCUNGSSSCFSLVPPMTIY 300

```


Db 241 SNTIDTSSQCIKRWTPPTRIYCLPSGIFVCGTSAVHCLNGSSSESMCFSLFVPPMTIY 300
 QY 301 TEODLYSVYISKPRNKRVPILPFVIGAGVLGALGICIGITSTQRYVYLSQELNDMER 360
 Db 301 TEODLYSVYISKPRNKRVPILPFVIGAGVLGALGICIGITSTQRYVYLSQELNDMER 360
 QY 361 VADSLVTLQDQNSLAAVVLAQNRALDILTAERGCTLGECECCYVNSGIVETKYE 420
 Db 361 VADSLVTLQDQNSLAAVVLAQNRALDILTAERGCTLGECECCYVNSGIVETKYE 420
 QY 421 TDRDQRRAREELRNTGPGWGLSQWMPWILPFLGPIALITLLIFGFCITNLLVNEVSSRI 480
 Db 421 TDRDQRRAREELRNTGPGWGLSQWMPWILPFLGPIALITLLIFGFCITNLLVNEVSSRI 480
 QY 481 EAVKLOMEPKMQ 492
 Db 481 EAVKLOMEPKMQ 492

RESULT 15

ABG51885
 ID ABG51885 standard; Peptide; 263 AA.

AC ABG51885;

DT 25-FEB-2003 (first entry)

DE Human liver peptide, SEQ ID No 30533.

KW Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;

KW hypercholesterolaemia; coronary heart disease.

OS Homo sapiens.

PN WO200157273-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00664.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DK;

DR WPI; 2001-488898/53.

PT Human genome-derived single exon nucleic acid probes useful for

PT analysing gene expression in human adult liver -

PS Claim 27; SEQ ID No 30533; 658bp; English.

CC The invention relates to a single exon nucleic acid probe (SENP) (I) for

CC measuring human gene expression in a sample derived from human adult

CC liver, comprising one of 13109 defined nucleotide sequences given in the

CC specification (or complements/ fragments). The probe hybridises at high

CC stringency to a nucleic acid molecule expressed in the human adult

CC liver. (I) may be used for predicting, measuring and displaying gene

CC expression in samples derived from human adult liver. The genes

CC identified may be involved in genetic liver diseases such as cirrhosis,

CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which

CC is associated with coronary heart disease. ABG47348-ABG59930 represent

CC human liver single exon encoded peptides of the invention.

CC Note: The sequence information for this patent does not appear in the

CC printed specification but was obtained in electronic format directly

CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 263 AA:
 Query Match 43.0%; Score 1234.5; DB 22; Length 263;
 Best local Similarity 89.4%; Pred. No. 1.1e-106;
 Matches 236; Conservative 11; Mismatches 16; Indels 1; Gaps 1;

QY 126 EKHVKEVISQLRVHGTSSPYKGLDLSKIHETLRTHTRLVSLFNTLTGLHVSANPTN 185
 Db 1 EKHKEVISQLRWVHSTPGPYRGDLDSKLHETLHTHTGLVSLFNTLTGLHVSANPTN 60
 QY 186 CWICLPLNRPYVSIPEPQWNNFSTEINTTSYLVGPLYSNLEITHTSNLTCVKSNTTY 245
 Db 61 CWMCLPLHFRPYISIPPEQWNNFSTEINTTSYLVGPLYSNLEITHTSNLTCVKSNTTY 119
 QY 246 TTNSOCIRWVTPPTQIVCLPSGIFVCGTSAVHCLNGSSSESMCFSLFVAPMPITYEODL 305
 Db 120 TANSOCIRWVTPPTQIVCLPSGIFVCGTSAVHCLNGSSSESMCFSLFVAPMPITYEODL 179
 QY 306 YSYVYSKPRNKRVPILPFVIGAGVLGALGICIGITSTQRYVYLSQELNDMERVADS 365
 Db 180 YNHVYIPKPRNKRVPILPFVIGAGVLGALGICIGITSTQRYVYLSQELNDMERVADS 239
 QY 366 VTLQDQNSLAAVVLAQNRALDIL 389
 Db 240 VTLQDQNSLAAVVLAQNRALDIL 263

Search completed: September 23, 2003, 14:54:47
 Job time : 88 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 23, 2003, 14:57:22 ; Search time 67 Seconds
(without alignments)
1195.592 Million cell updates/sec

Title: US-10-069-883-1

Perfect score: 2871
Sequence: 1 MALPYHIFLFTVLPSTFLT.....PPEISAQPLIRPNSAGSS 538

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 556269 seqs, 148893369 residues

Total number of hits satisfying chosen parameters: 556269

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: PublishedApplications_AA.*
2: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/PCR_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/PCR_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2868	99.9	538	US-09-902-535-2	Sequence 2, Appl
2	2868	99.9	538	US-10-114-893-135	Sequence 135, App
3	2868	99.9	538	US-10-016-249-4	Sequence 4, Appl
4	2314	80.6	493	US-08-979-847-106	Sequence 106, App
5	1234.5	43.0	263	US-08-864-761-37656	Sequence 37656, A
6	672	23.4	162	US-08-979-847-109	Sequence 109, App
7	563.5	19.6	576	US-09-793-451-4	Sequence 4, Appl
8	530.5	18.5	466	US-10-001-835-168	Sequence 168, App
9	484	16.9	433	US-08-979-847-81	Sequence 81, Appl
10	464	16.2	253	US-09-764-891-2970	Sequence 2970, App
11	411.5	14.3	638	US-09-851-859A-6	Sequence 6, Appl
12	403.5	14.1	660	US-09-851-859A-5	Sequence 5, Appl
13	390.5	13.6	657	US-09-851-859A-4	Sequence 4, Appl
14	386.5	13.5	656	US-09-851-859A-3	Sequence 3, Appl
15	370	12.9	67	US-10-040-916-49	Sequence 49, Appl

16	287	10.0	77	US-08-979-847-103	Sequence 103, App
17	285.5	9.9	227	US-09-739-254-126	Sequence 126, App
18	285.5	9.9	227	US-09-904-615-126	Sequence 126, App
19	285.5	9.9	227	US-10-055-098-126	Sequence 126, App
20	285.5	9.9	227	US-10-054-988-126	Sequence 126, App
21	251.5	8.8	131	US-09-864-761-47888	Sequence 47888, A
22	217	7.6	63	US-09-864-761-40868	Sequence 40868, A
23	204	7.1	63	US-09-864-761-37442	Sequence 37442, A
24	175	6.1	50	US-09-864-761-35377	Sequence 35377, A
25	171.5	6.0	53	US-09-793-451-2	Sequence 2, Appl
26	167.5	5.8	563	US-09-789-561-83	Sequence 83, Appl
27	147	5.1	36	US-09-864-761-40624	Sequence 40624, A
28	146.5	5.1	681	US-10-066-506A-14	Sequence 14, Appl
29	145	5.1	681	US-10-066-506A-6	Sequence 6, Appl
30	143	5.0	681	US-10-267-322-2	Sequence 2, Appl
31	141	4.9	681	US-10-066-506A-10	Sequence 10, Appl
32	136.5	4.8	681	US-10-066-506A-8	Sequence 8, Appl
33	131	4.6	747	US-10-066-506A-2	Sequence 2, Appl
34	126	4.4	54	US-09-867-550-1804	Sequence 1804, App
35	124.5	4.3	610	US-10-066-506A-4	Sequence 4, Appl
36	121	4.2	596	US-09-862-021-21	Sequence 21, Appl
37	109	3.8	27	US-09-739-254-128	Sequence 128, App
38	109	3.8	27	US-09-904-615-128	Sequence 128, App
39	109	3.8	27	US-10-055-098-128	Sequence 128, App
40	109	3.8	27	US-10-054-988-128	Sequence 128, App
41	105.5	3.7	676	US-10-066-506A-12	Sequence 12, Appl
42	104.5	3.6	676	US-09-337-946A-17	Sequence 17, Appl
43	104	3.6	854	US-10-369-294-16	Sequence 16, Appl
44	99.5	3.5	862	US-10-369-294-15	Sequence 15, Appl
45	98.5	3.4	641	US-10-142-515-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-09-902-535-2
; Sequence 2, Application US/09902535
; Patent No. US20020102530A1
; GENERAL INFORMATION:
; APPLICANT: Keltch, Jr., James C.
; APPLICANT: McCoy, John M.
; APPLICANT: M, Sha
; TITLE OF INVENTION: Methods and compositions for diagnosing
; TITLE OF INVENTION: and treating preclampsia and gestational trophoblast
; FILE OF INVENTION: disorders
; FILE REFERENCE: GIN-600684
; CURRENT APPLICATION NUMBER: US/09/902,535
; PRIOR APPLICATION NUMBER: 2001-07-09
; PRIOR FILING DATE: 2000-07-06
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 538
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-902-535-2

Query Match 99.9%; Score 2868; DB 10; Length 538;
Best local Similarity 99.8%; Pred. No. 2e-267;
Matches 537; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALPYHIFLFTVLPSTFLTAPPCRCMTSSSYOEFILRMORPGIDAPSTRSLKGP 60
|||||
DB 1 MALPYHIFLFTVLPSTFLTAPPCRCMTSSSYOEFILRMORPGIDAPSTRSLKGP 60
|||||
QY 61 TFAHTHMRNCYHSATLGMHANTHYWTKMIMPSCGGLGYVCWTYTTOTGMSDGGCV 120
|||||
DB 61 TFAHTHMRNCYHSATLGMHANTHYWTKMIMPSCGGLGYVCWTYTTOTGMSDGGCV 120
|||||
QY 121 ODOARKHKEYVISOLTRVHGTSPPYKGLDKLHETLRTPHRLVSLFWTTLTGHEVGA 180
|||||

Db 121 ODOAREKHVKEVVISQLTRVHGTSPPYKGLDLSKIHETLRTHRLVSLFNTTLTGJLHEVSA 180
QY 181 ONPTNCWICLPINFRPVYIPVPEOMNFEINTEINTSVLGVPLVSNLEITHSNLTGVKF 240
Db 181 ONPTNCWICLPINFRPVYIPVPEOMNFEINTEINTSVLGVPLVSNLEITHSNLTGVKF 240
QY 241 SNTYTTNSQCIIRWTPPTQIVCLPSGIFFCGTSAYACLNSSSESMCFSLFVLPMTIY 300
Db 241 SNTYTTNSQCIIRWTPPTQIVCLPSGIFFCGTSAYACLNSSSESMCFSLFVLPMTIY 300
QY 301 TEODLYSVISKPRNKRVPILPFVYGAVLGCALGTGIGITTSQFYKLSQELNGDMER 360
Db 301 TEODLYSVISKPRNKRVPILPFVYGAVLGCALGTGIGITTSQFYKLSQELNGDMER 360
QY 361 VADSLVTLQDOLNSLAAYVLONNRRLDLTLTAERGTCFLFEBCCYVYNOGSIYTEKYE 420
Db 361 VADSLVTLQDOLNSLAAYVLONNRRLDLTLTAERGTCFLFEBCCYVYNOGSIYTEKYE 420
QY 421 IRDRIOARRAEELRNTGPMGLLSQWMPWILPFLGPLAAIILLLLFGPCIFNLLVNFVSRI 480
Db 421 IRDRIOARRAEELRNTGPMGLLSQWMPWILPFLGPLAAIILLLLFGPCIFNLLVNFVSRI 480
QY 481 EAVKLQMEPKMQSKTKIYRRPLDRPASPRSDVNDIKGTPPEISAAOPLLRPNASGSS 538
Db 481 EAVKLQMEPKMQSKTKIYRRPLDRPASPRSDVNDIKGTPPEISAAOPLLRPNASGSS 538

RESULT 2

US-10-114-893-135
: Sequence 135, Application US/10114893
: Publication No. US20020193567A1
: GENERAL INFORMATION:
: APPLICANT: Jacobs, Kenneth
: APPLICANT: McCoy, John M.
: APPLICANT: Lavalley, Edward R.
: APPLICANT: Collins-Racie, Lisa A.
: APPLICANT: Evans, Cheryl
: APPLICANT: Merberg, David
: APPLICANT: Treacy, Maurice
: APPLICANT: Bowman, Michael R.
: APPLICANT: Paulding, Vikki
: APPLICANT: Carlin-Duckett, McKeough
: APPLICANT: Kelleher, Kerry S.
: APPLICANT: Genetics Institute, Inc.
: TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
: FILE REFERENCE: GI 6000-10A
: CURRENT APPLICATION NUMBER: US/10/114, 893
: CURRENT FILING DATE: 2002-04-02
: EARLIER APPLICATION NUMBER: 09/413, 232
: EARLIER FILING DATE: 1999-10-06
: NUMBER OF SEQ ID NOS: 321
: SOFTWARE: Patent Ver. 2.0
: SEQ ID NO 135
: LENGTH: 538
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-114-893-135

Query Match 99.9%; Score 2868; DB 14; Length 538;
Best Local Similarity 99.8%; Pred. No. 2e-267;
Matches 537; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALPYHIFLFTVLLPSFTLTAPPCRCMTSSPYOEFLMRQRCNIDAPYSRSLSKGTP 60
Db 1 MALPYHIFLFTVLLPSFTLTAPPCRCMTSSPYOEFLMRQRCNIDAPYSRSLSKGTP 60
QY 61 TFTAHTHPRNCYSATLCHMANHYWTKMINSPCGGLGVTCWMTYFTTQGMSDGGV 120
Db 61 TFTAHTHPRNCYSATLCHMANHYWTKMINSPCGGLGVTCWMTYFTTQGMSDGGV 120
QY 121 ODOAREKHVKEVVISQLTRVHGTSPPYKGLDLSKIHETLRTHRLVSLFNTTLTGJLHEVSA 180
Db 121 ODOAREKHVKEVVISQLTRVHGTSPPYKGLDLSKIHETLRTHRLVSLFNTTLTGJLHEVSA 180

QY 181 ONPTNCWICLPINFRPVYIPVPEOMNFEINTEINTSVLGVPLVSNLEITHSNLTGVKF 240
Db 181 ONPTNCWICLPINFRPVYIPVPEOMNFEINTEINTSVLGVPLVSNLEITHSNLTGVKF 240
QY 241 SNTYTTNSQCIIRWTPPTQIVCLPSGIFFCGTSAYACLNSSSESMCFSLFVLPMTIY 300
Db 241 SNTYTTNSQCIIRWTPPTQIVCLPSGIFFCGTSAYACLNSSSESMCFSLFVLPMTIY 300
QY 301 TEODLYSVISKPRNKRVPILPFVYGAVLGCALGTGIGITTSQFYKLSQELNGDMER 360
Db 301 TEODLYSVISKPRNKRVPILPFVYGAVLGCALGTGIGITTSQFYKLSQELNGDMER 360
QY 361 VADSLVTLQDOLNSLAAYVLONNRRLDLTLTAERGTCFLFEBCCYVYNOGSIYTEKYE 420
Db 361 VADSLVTLQDOLNSLAAYVLONNRRLDLTLTAERGTCFLFEBCCYVYNOGSIYTEKYE 420
QY 421 IRDRIOARRAEELRNTGPMGLLSQWMPWILPFLGPLAAIILLLLFGPCIFNLLVNFVSRI 480
Db 421 IRDRIOARRAEELRNTGPMGLLSQWMPWILPFLGPLAAIILLLLFGPCIFNLLVNFVSRI 480
QY 481 EAVKLQMEPKMQSKTKIYRRPLDRPASPRSDVNDIKGTPPEISAAOPLLRPNASGSS 538
Db 481 EAVKLQMEPKMQSKTKIYRRPLDRPASPRSDVNDIKGTPPEISAAOPLLRPNASGSS 538

RESULT 3

US-10-016-249-4
: Sequence 4, Application US/10016249
: Publication No. US20030100053A1
: GENERAL INFORMATION:
: APPLICANT: Jacobs, Kenneth
: APPLICANT: McCoy, John M.
: APPLICANT: Lavalley, Edward R.
: APPLICANT: Collins-Racie, Lisa A.
: APPLICANT: Evans, Cheryl
: APPLICANT: Merberg, David
: APPLICANT: Treacy, Maurice
: APPLICANT: M. Sha
: APPLICANT: Genetics Institute, Inc.
: TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
: FILE REFERENCE: 6006B, A172A
: CURRENT APPLICATION NUMBER: US/10/016, 249
: CURRENT FILING DATE: 2001-10-30
: PRIOR APPLICATION NUMBER: US/09/175, 928
: PRIOR FILING DATE: 1998-10-20
: NUMBER OF SEQ ID NOS: 62
: SOFTWARE: Patent Ver. 2.0
: SEQ ID NO 4
: LENGTH: 538
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-016-249-4

Query Match 99.9%; Score 2868; DB 15; Length 538;
Best Local Similarity 99.8%; Pred. No. 2e-267;
Matches 537; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALPYHIFLFTVLLPSFTLTAPPCRCMTSSPYOEFLMRQRCNIDAPYSRSLSKGTP 60
Db 1 MALPYHIFLFTVLLPSFTLTAPPCRCMTSSPYOEFLMRQRCNIDAPYSRSLSKGTP 60
QY 61 TFTAHTHPRNCYSATLCHMANHYWTKMINSPCGGLGVTCWMTYFTTQGMSDGGV 120
Db 61 TFTAHTHPRNCYSATLCHMANHYWTKMINSPCGGLGVTCWMTYFTTQGMSDGGV 120
QY 121 ODOAREKHVKEVVISQLTRVHGTSPPYKGLDLSKIHETLRTHRLVSLFNTTLTGJLHEVSA 180
Db 121 ODOAREKHVKEVVISQLTRVHGTSPPYKGLDLSKIHETLRTHRLVSLFNTTLTGJLHEVSA 180
QY 181 ONPTNCWICLPINFRPVYIPVPEOMNFEINTEINTSVLGVPLVSNLEITHSNLTGVKF 240
Db 181 ONPTNCWICLPINFRPVYIPVPEOMNFEINTEINTSVLGVPLVSNLEITHSNLTGVKF 240

RESULT 6
 US-08-979-847-109
 : Sequence 109, Application US/08979847
 : Publication No. US2003003664A1
 :
 : GENERAL INFORMATION:
 :
 : APPLICANT: PERRON, HERVE
 : APPLICANT: BESEME, FREDERIC
 : APPLICANT: BEDIN, FREDERIC
 : APPLICANT: PARANHOS-BACCALA, GLAUCIA

```

: RESULT 7
: US-09-793-451-4
:
: Sequence 4, Application US/09793451
: Publication NO. US20030157597A1
: GENERAL INFORMATION:
:
: APPLICANT: Arthur B. Raitano
:
: APPLICANT: Daniel E.H. Afar
:
: APPLICANT: Gazelle S. Rastegar
:
: APPLICANT: Steve Chappell Mitchell
:
: APPLICANT: Renee S. Hubert
:
: APPLICANT: pia M. Challita-Eid
:
: APPLICANT: Mary Faris
:
: APPLICANT: Aya Jakobovits
:
: TITLE OF INVENTION: 103P2D6: TISSUE SPECIFIC PROTEIN HIGHLY
: TITLE OF INVENTION: EXPRESSED IN VARIOUS CANCERS

```

Query Match	19.68	Score 563.5	DB 12	Length 576
Best Local Similarity	27.18	Pred. No. 3.2e-45		
Matches 152; Conservative	66	Mismatches 122;	Indels 221;	Gaps 17

RESULT 8
US-10-001-835-16E

```

CURRENT APPLICATION NUMBER: US/10/001,835
CURRENT FILING DATE: 2001-11-20
PRIOR APPLICATION NUMBER: 60/249,997
PRIORITY FILING DATE: 2000-11-20
NUMBER OF SEQ ID NOS: 228
SOFTWARE: PatentIn version 3.1
SEQ ID NO 168
LENGTH: 469
TYPE: PRT
ORGANISM: Homo sapien
US-10-001-835-168

Query Match      18.5%; Score 530.5; DB 14; Length 469;
Best Local Similarity   33.6%; Pred. No. 3-Se-42;
Matches 144; Conservative    64; Mismatches 152; Indels     69; Gaps    16;

OY      LTRVHGSSPKKGDLSKLHETLR-----TH--TRLVSLFNTTLTGHEVSADQPTN----185
DB       ::::|::||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY      ISKRPRGIPGLGNLSPRCSFTLHRBPTHHTTETIGAFOLAHITDPSIWTDXLKINSN 112
DB       :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY      186 -C-----WICLPINF---RPVVISPYPEOMNFSTEI-----NTTS 217
DB       |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY      113 YCLGRHLPSISLHPWLPPCSSDSPRRSRLILSRKNNSERLVDTQRELIHHENRTS 172
DB       ||::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY      218 VLVRPLVSLNEITHITSULTCVKKFSNTTYTTNSOCIRAV-----TPPTQLV----CLPSG 267
DB       :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY      173 -----PSIQPLHQSDAPL---TAASLAGSLGIIVODTPSTPLFLTLHIQCCTOG 221
DB       :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY      268 IFEVCGSATRCNLGSSESMCFSELVPMTITY-TEDLYSVYSKPNNRV-PILPFVI 325
DB       |:|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY      222 LFELCGSSTVMCLPANMTCITVLVELFKIQANGTELEYPMLTFPRKARVIPLLPMAY 281
DB       :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY      326 GAGVGAG--ALGTGIGITTSTOFYYKLSOELNGDMERVADSIVLTDDQSILAAYVLON 382
DB       |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY      282 GLGISASTIALGTGIAGISTVTFFRSLSNDIFSASTIDISQTLSYLQAQVDLAAYVLON 341
DB       :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY      383 RRAIDLIAEKGCTCLFGEECCYYNVOSGITVEKKVKIEDRIQRAREBLRNMG--PWGL 440
DB       |||||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY      342 RRGLDLTAEGGGCIPIINEGCCFYLNQSGIYVDNIKLDRAOKLANQASNABEPWA - 400
DB       :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY      441 LSQMNPILLPLPGPLAIIIIIIIIFGCIFIMLVNFVSRSIEAKLIKMEPMQSKTKRYR 500
DB       |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY      401 LSNRMWSVLPLSLPLIPLFIPLLFAPCIFCVLOSFIQNRIQAIINHNSIAMFILTIPOXH 460
DB       :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

OY      501 PL--DRPAS 507
DB       |||::|
OY      461 PLPDLPESA 469

RESULT 9
US-08-979-847-81
Sequence 81, Application US/08979847
Publication No. US20030039664A1
GENERAL INFORMATION:
APPLICANT: PERRON, HERVE
APPLICANT: BESEMME, FREDERIC
APPLICANT: BEDIN, FREDERIC
APPLICANT: PARANHOS-BACCALA, GLAUCIA
APPLICANT: KOMURIAN-PRODEL, FLORENCE
APPLICANT: JOLIVERET-RETNAUD, COLETTE
APPLICANT: MANDEBRAND, BERNARD
APPLICANT: GARSON, JEREMY
APPLICANT: TUKE, PHILIP
TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
TITLE OF INVENTION: ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYL
NUMBER OF SEQUENCES: 210
CORRESPONDENCE ADDRESSES:
ADDRESSEE: OLIFF & HERRIDGE, PLC
STREET: P.O. BOX 19928
CITY: ALEXANDRIA
STATE: VA
COUNTRY: USA
```

Query Match	16.9%	Score 484	DB 8	Length 433
Best Local Similarly	53.0%	Pred. No. 9.5e-38		
Matches 97	Conservative 27	Mismatches 41	Indels 18	Gaps 4

RESULT 10
US-09-764-891-2970

```

: Sequence 2970, Application US/09764891
: Publication No. US20030077808A1
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
: FILE REFERENCE: PC006
: CURRENT APPLICATION NUMBER: US/09/764,891
: CURRENT FILING DATE: 2001-01-17
: Prior application data removed - consult PALM or file wrapper
: NUMBER OF SEQ ID NOS: 10231
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 2970
: LENGTH: 253
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: SITE
: LOCATION: (162)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
: NAME/KEY: SITE
: LOCATION: (195)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
: NAME/KEY: SITE
: LOCATION: (206)

```

Query Match	16.28;	Score 464;	DB 11;	Length 253;
Best Local Similarity	40.68;	Pred. No. 3.5e-36;		
Matches 93;	Conservative 49;	Mismatches 59;	Indels 28;	Gaps 5

RESULT 11
US-09-851-859A-6

```

? Sequence 6 Application US//09851859A
? Patent No. US20020065407A1
? GENERAL INFORMATION:
? APPLICANT: Banerjee, Papia T.
? APPLICANT: Patience, Clive
? TITLE OF INVENTION: Molecular Sequence of Swine Retrovirus and Methods of
? TITLE OF INVENTION: Use
? FILE REFERENCE: 61750-321
? CURRENT APPLICATION NUMBER: US/09/851,859A
? PRIOR FILING DATE: 1999-08-18
? PRIOR FILING DATE: 1998-08-18
? PRIOR APPLICATION NUMBER: US 09/376781
? PRIOR FILING DATE: 2000-08-18
? NUMBER OF SEQ ID NOS: 33
? SOFTWARE: PatentIn Ver. 2.1
? SEQ ID NO 6
? LENGTH: 638
? TYPE: PRT
? ORGANISM: Artificial Sequence
? FEATURE:
? OTHER INFORMATION: Description of Artificial Sequence: PERV-C
? OTHER INFORMATION: polypeptide sequence taken from GenBank Accession
? OTHER INFORMATION: No. US20020065407A1 AF038600 for comparison.
? US-09-851-859A-6

```

	14.38:	Score	411.5:	DB:9:	Length	638;
Query Match						
Best Local Similarity	24.78;		Pred. No.	1.7e-30;		
Matches 136;	Conservative	76;	Mismatches	208;	Indels	131; Gaps 19
QY	58	GPEPTAHTHMPRNCYHSATLCMHANTHW-----TGKMINPSCPGGLGVTCW-	106			
	: :	:	:			
Db	115	GPNNKHGHNPRDFCKOMNCVTSNDGCKWKPSTSODRVSEFVSYYNTTSSGGPNYLTWI	174			
QY	107	-----TYFTQTGMSPD-----GGGYDQAENEKRYKEVIS	124			
	:	:	:			
Db	175	RIGSPKCPSPDLDYLKISTFEKGKQENILKWNCMSMGMYVYGSGKPGSILTIRLKTIN	234			
QY	135	QL-----TRVHGTSPPY-KGLDISKLHEETLRTHRLVSLFWTTTLTGL	175			
	:	:	:			
Db	235	QLBPMAIGNVTVLGTGRPETGQPPSSNITSGSPTESSSTTKMKAKLFELIGARFOAL	294			
QY	176	HESVAQNPNRCWCLEPLNRPVYSIPVEQMNPFSTELNTITSYLGLPVLSMLETHHSNL	235			
	:	:	:			
Db	295	NSTTPPATSYSCWICLASG-----PPLYEGAMRR-----GKNVYTKHRD	333			

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - protein search, using sw model

Run on: September 23, 2003, 14:54:52 : Search time 403 Seconds
(without alignments)
1214.730 Million cell updates/sec

Title: US-10-069-883-1

Perfect score: 2871
Sequence: 1 MALPYHIFLFTLPSPFTLT.....PPEISAAQPLRPMSAGSS 538

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 5728757 seqs, 909918778 residues

Total number of hits satisfying chosen parameters: 5728757

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Pending Patents_AA_Main:*
1: /cgn2_6/ptodata/1/paa/PCFUS_COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US08_COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US081_COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US082_COMB.pep:*
7: /cgn2_6/ptodata/1/paa/US083_COMB.pep:*
8: /cgn2_6/ptodata/1/paa/US084_COMB.pep:*
9: /cgn2_6/ptodata/1/paa/US085_COMB.pep:*
10: /cgn2_6/ptodata/1/paa/US086_COMB.pep:*
11: /cgn2_6/ptodata/1/paa/US087_COMB.pep:*
12: /cgn2_6/ptodata/1/paa/US088_COMB.pep:*
13: /cgn2_6/ptodata/1/paa/US089_COMB.pep:*
14: /cgn2_6/ptodata/1/paa/US090_COMB.pep:*
15: /cgn2_6/ptodata/1/paa/US091_COMB.pep:*
16: /cgn2_6/ptodata/1/paa/US092_COMB.pep:*
17: /cgn2_6/ptodata/1/paa/US093_COMB.pep:*
18: /cgn2_6/ptodata/1/paa/US094_COMB.pep:*
19: /cgn2_6/ptodata/1/paa/US095_COMB.pep:*
20: /cgn2_6/ptodata/1/paa/US096_COMB.pep:*
21: /cgn2_6/ptodata/1/paa/US097A_COMB.pep:*
22: /cgn2_6/ptodata/1/paa/US097B_COMB.pep:*
23: /cgn2_6/ptodata/1/paa/US098_COMB.pep:*
24: /cgn2_6/ptodata/1/paa/US099A_COMB.pep:*
25: /cgn2_6/ptodata/1/paa/US099B_COMB.pep:*
26: /cgn2_6/ptodata/1/paa/US100_COMB.pep:*
27: /cgn2_6/ptodata/1/paa/US101_COMB.pep:*
28: /cgn2_6/ptodata/1/paa/US102_COMB.pep:*
29: /cgn2_6/ptodata/1/paa/US103_COMB.pep:*
30: /cgn2_6/ptodata/1/paa/US104_COMB.pep:*
31: /cgn2_6/ptodata/1/paa/US106_COMB.pep:*
32: /cgn2_6/ptodata/1/paa/US60_COMB.pep:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2871	100.0	538	22	US-09-791-537-133031
					Sequence 133031,

2	2871	100.0	538	26	US-10-069-883-1	Sequence 1, Appl1
3	2871	100.0	538	27	US-10-133-036-2	Sequence 2, Appl1
4	2871	100.0	540	21	US-09-719-554-26	Sequence 26, Appl1
5	2871	100.0	540	24	US-09-949-016-11511	Sequence 11511, A
6	2871	100.0	685	21	US-09-719-554-35	Sequence 35, Appl1
7	2868	99.9	538	1	PCT-US98-22034-4	Sequence 4, Appl1
8	2868	99.9	538	1	PCT-US98-24614-4	Sequence 4, Appl1
9	2868	99.9	538	1	PCT-US99-10915-4	Sequence 4, Appl1
10	2868	99.9	538	13	US-08-976-110-4	Sequence 4, Appl1
11	2868	99.9	538	14	US-09-080-218-4	Sequence 4, Appl1
12	2868	99.9	538	18	US-09-413-233-135	Sequence 135, Appl1
13	2868	99.9	538	22	US-09-791-537-31643	Sequence 31643, A
14	2868	99.9	538	24	US-09-902-535-2	Sequence 2, Appl1
15	2868	99.9	538	26	US-10-016-249-4	Sequence 4, Appl1
16	2868	99.9	538	27	US-10-114-893-135	Sequence 135, Appl1
17	2853	99.4	538	26	US-10-069-883-57	Sequence 57, Appl1
18	2849	99.2	538	26	US-10-069-883-61	Sequence 61, Appl1
19	2846	99.1	538	26	US-10-069-883-59	Sequence 59, Appl1
20	2842	99.0	538	27	US-10-133-036-21	Sequence 21, Appl1
21	2839	98.9	538	26	US-10-069-883-55	Sequence 55, Appl1
22	2791.5	97.2	544	1	PCT-US02-25485-2537	Sequence 2537, Ap
23	2791	97.2	538	18	US-09-446-024-333	Sequence 33, Appl1
24	2791	97.2	538	18	US-09-446-024A-33	Sequence 33, Appl1
25	2791	97.2	538	27	US-10-133-036-20	Sequence 20, Appl1
26	2558	89.1	542	30	US-10-416-642-2	Sequence 2, Appl1
27	2490	86.7	542	24	US-09-936-835-2	Sequence 2, Appl1
28	2336	81.4	503	23	US-09-890-340-20	Sequence 20, Appl1
29	2336	81.4	503	31	US-10-637-565-20	Sequence 20, Appl1
30	2314	80.6	493	17	US-09-319-156A-10	Sequence 10, Appl1
31	2314	80.6	493	17	US-10-114-104-106	Sequence 106, App
32	2314	80.6	493	27	US-10-114-104A-106	Sequence 106, App
33	2314	80.6	493	30	US-10-446-730-106	Sequence 106, App
34	1249	43.5	256	32	US-60-453-050-10770	Sequence 10770, A
35	1249	43.5	256	32	US-60-453-135-10770	Sequence 10770, A
36	1249	43.5	256	32	US-60-466-412-10770	Sequence 10770, A
37	1234.5	43.0	263	1	PCT-US01-00663-30788	Sequence 30788, A
38	1234.5	43.0	263	23	US-09-864-761-37656	Sequence 37656, A
39	1234.5	43.0	263	27	US-10-182-995-29875	Sequence 29875, A
40	1234.5	43.0	263	27	US-10-182-995-24128	Sequence 24128, A
41	1234.5	43.0	263	27	US-10-182-995-22834	Sequence 22834, A
42	1234.5	43.0	263	27	US-10-182-998-14388	Sequence 14388, A
43	1234.5	43.0	263	28	US-10-203-134-30491	Sequence 30491, A
44	1234.5	43.0	263	28	US-10-203-135-29484	Sequence 29484, A
45	1234.5	43.0	263	28	US-10-203-136-30533	Sequence 30533, A

ALIGNMENTS

RESULT 1
US-09-791-537-133031
; Sequence 133031, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY M-
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 133031
; LENGTH: 538
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-791-537-133031

Query Match 100.0%; Score 2871; DB 22; Length 538;
Best Local Similarity 100.0%; Pred. No. 9, 1e-275;
Matches 538; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MALPYHIFLFTVLLPSTLTAPPCCRCMTSSSPYOEFLMRQPGNIDAPSYRSLSKGP 60
DB 1 MALPYHIFLFTVLLPSTLTAPPCCRCMTSSSPYOEFLMRQPGNIDAPSYRSLSKGP 60
QY 61 TTTAHTHMPRNCYHSATLCMHANTHYWGKMINPSCPGGLGYTCVCTTFTOTGSDGCV 120
DB 61 TTTAHTHMPRNCYHSATLCMHANTHYWGKMINPSCPGGLGYTCVCTTFTOTGSDGCV 120
QY 121 ODOAREKHVKEVYSQLTRVHGTSSPYKGLDLSKLHETLRTHTRLVSLFNTTLTGLHEVSA 180
DB 121 ODOAREKHVKEVYSQLTRVHGTSSPYKGLDLSKLHETLRTHTRLVSLFNTTLTGLHEVSA 180
QY 181 OQOAREKHVKEVYSQLTRVHGTSSPYKGLDLSKLHETLRTHTRLVSLFNTTLTGLHEVSA 180
DB 181 OQOAREKHVKEVYSQLTRVHGTSSPYKGLDLSKLHETLRTHTRLVSLFNTTLTGLHEVSA 180
QY 241 SNTTNTTSOCIRMTWTPPTQIVCLPSGLFVCGTSAYKCLNGSSSMCFSLVLPMTIY 300
DB 241 SNTTNTTSOCIRMTWTPPTQIVCLPSGLFVCGTSAYKCLNGSSSMCFSLVLPMTIY 300
QY 301 TEODLYSYVSKPRNKRVPILPFVIGAVLGALGTGIGITTSFTQFYKLSQELNGDME 360
DB 301 TEODLYSYVSKPRNKRVPILPFVIGAVLGALGTGIGITTSFTQFYKLSQELNGDME 360
QY 361 VADSLVLTQDOLNSLAAYVLONRALDLTLTAERGTCFLGEECCYVYNOGSIYTEKYE 420
DB 361 VADSLVLTQDOLNSLAAYVLONRALDLTLTAERGTCFLGEECCYVYNOGSIYTEKYE 420
QY 421 IRDRIORAEELRNTGPMGLLSOMMPWILPFLGPIAAIITLLRPCLFNLVNVSSRI 480
DB 421 IRDRIORAEELRNTGPMGLLSOMMPWILPFLGPIAAIITLLRPCLFNLVNVSSRI 480
QY 481 EAVKLQMPKMQSKTKIYRRPLDRPASPRSDVNDIKGTPPEISAAQPLRPNSAGSS 538
DB 481 EAVKLQMPKMQSKTKIYRRPLDRPASPRSDVNDIKGTPPEISAAQPLRPNSAGSS 538

```

RESULT 2

```

US-10-069-883-1
; Sequence 1, Application US/10069883
; GENERAL INFORMATION:
; APPLICANT: MALLET, Francois
; APPLICANT: COSSET, Francois-Joic
; APPLICANT: BLOND, Jean-Luc
; APPLICANT: LAVILLETTE, Dimalri
; APPLICANT: BOUON, Olivier
; APPLICANT: RUGIERI, Alessia
; TITLE OF INVENTION: Method for detecting the expression of an envelope protein of a
; TITLE OF INVENTION: endogenous retrovirus and uses of a gene coding for said protein
; FILE REFERENCE: 112062
; CURRENT APPLICATION NUMBER: US/10/069, 883
; CURRENT FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: PCT/FR00/02429
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: FR-9911141
; PRIOR FILING DATE: 1999-09-01
; PRIOR APPLICATION NUMBER: FR-9911793
; PRIOR FILING DATE: 1999-09-15
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 538
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-069-883-1

```

```

Query Match 100.0%; Score 2871; DB 26; Length 538;
Best Local Similarity 100.0%; Pred. No. 9,1e-275;
Matches 538; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MALPYHIFLFTVLLPSTLTAPPCCRCMTSSSPYOEFLMRQPGNIDAPSYRSLSKGP 60
DB 1 MALPYHIFLFTVLLPSTLTAPPCCRCMTSSSPYOEFLMRQPGNIDAPSYRSLSKGP 60

```

```

QY 61 TTTAHTHMPRNCYHSATLCMHANTHYWGKMINPSCPGGLGYTCVCTTFTOTGSDGCV 120
DB 61 TTTAHTHMPRNCYHSATLCMHANTHYWGKMINPSCPGGLGYTCVCTTFTOTGSDGCV 120
QY 121 ODOAREKHVKEVYSQLTRVHGTSSPYKGLDLSKLHETLRTHTRLVSLFNTTLTGLHEVSA 180
DB 121 ODOAREKHVKEVYSQLTRVHGTSSPYKGLDLSKLHETLRTHTRLVSLFNTTLTGLHEVSA 180
QY 181 OQOAREKHVKEVYSQLTRVHGTSSPYKGLDLSKLHETLRTHTRLVSLFNTTLTGLHEVSA 180
DB 181 OQOAREKHVKEVYSQLTRVHGTSSPYKGLDLSKLHETLRTHTRLVSLFNTTLTGLHEVSA 180
QY 241 SNTTNTTSOCIRMTWTPPTQIVCLPSGLFVCGTSAYKCLNGSSSMCFSLVLPMTIY 300
DB 241 SNTTNTTSOCIRMTWTPPTQIVCLPSGLFVCGTSAYKCLNGSSSMCFSLVLPMTIY 300
QY 301 TEODLYSYVSKPRNKRVPILPFVIGAVLGALGTGIGITTSFTQFYKLSQELNGDME 360
DB 301 TEODLYSYVSKPRNKRVPILPFVIGAVLGALGTGIGITTSFTQFYKLSQELNGDME 360
QY 361 VADSLVLTQDOLNSLAAYVLONRALDLTLTAERGTCFLGEECCYVYNOGSIYTEKYE 420
DB 361 VADSLVLTQDOLNSLAAYVLONRALDLTLTAERGTCFLGEECCYVYNOGSIYTEKYE 420
QY 421 IRDRIORAEELRNTGPMGLLSOMMPWILPFLGPIAAIITLLRPCLFNLVNVSSRI 480
DB 421 IRDRIORAEELRNTGPMGLLSOMMPWILPFLGPIAAIITLLRPCLFNLVNVSSRI 480
QY 481 EAVKLQMPKMQSKTKIYRRPLDRPASPRSDVNDIKGTPPEISAAQPLRPNSAGSS 538
DB 481 EAVKLQMPKMQSKTKIYRRPLDRPASPRSDVNDIKGTPPEISAAQPLRPNSAGSS 538

```

RESULT 3

```

US-10-133-036-2
; Sequence 2, Application US/10133036
; GENERAL INFORMATION:
; APPLICANT: Conrad, Bernard
; APPLICANT: Mach, Bernard
; TITLE OF INVENTION: Multiple Sclerosis-Related Superantigen
; FILE REFERENCE: 23135-507
; CURRENT APPLICATION NUMBER: US/10/133, 036
; CURRENT FILING DATE: 2002-04-26
; PRIOR APPLICATION NUMBER: PCT/EP00/10659
; PRIOR FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 538
; TYPE: PRT
; ORGANISM: Human endogenous retrovirus
US-10-133-036-2

```

```

Query Match 100.0%; Score 2871; DB 27; Length 538;
Best Local Similarity 100.0%; Pred. No. 9,1e-275;
Matches 538; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MALPYHIFLFTVLLPSTLTAPPCCRCMTSSSPYOEFLMRQPGNIDAPSYRSLSKGP 60
DB 1 MALPYHIFLFTVLLPSTLTAPPCCRCMTSSSPYOEFLMRQPGNIDAPSYRSLSKGP 60
QY 61 TTTAHTHMPRNCYHSATLCMHANTHYWGKMINPSCPGGLGYTCVCTTFTOTGSDGCV 120
DB 61 TTTAHTHMPRNCYHSATLCMHANTHYWGKMINPSCPGGLGYTCVCTTFTOTGSDGCV 120
QY 121 ODOAREKHVKEVYSQLTRVHGTSSPYKGLDLSKLHETLRTHTRLVSLFNTTLTGLHEVSA 180
DB 121 ODOAREKHVKEVYSQLTRVHGTSSPYKGLDLSKLHETLRTHTRLVSLFNTTLTGLHEVSA 180
QY 181 OQOAREKHVKEVYSQLTRVHGTSSPYKGLDLSKLHETLRTHTRLVSLFNTTLTGLHEVSA 180
DB 181 OQOAREKHVKEVYSQLTRVHGTSSPYKGLDLSKLHETLRTHTRLVSLFNTTLTGLHEVSA 180
QY 241 SNTTNTTSOCIRMTWTPPTQIVCLPSGLFVCGTSAYKCLNGSSSMCFSLVLPMTIY 300
DB 241 SNTTNTTSOCIRMTWTPPTQIVCLPSGLFVCGTSAYKCLNGSSSMCFSLVLPMTIY 300
QY 301 TEODLYSYVSKPRNKRVPILPFVIGAVLGALGTGIGITTSFTQFYKLSQELNGDME 360
DB 301 TEODLYSYVSKPRNKRVPILPFVIGAVLGALGTGIGITTSFTQFYKLSQELNGDME 360
QY 361 VADSLVLTQDOLNSLAAYVLONRALDLTLTAERGTCFLGEECCYVYNOGSIYTEKYE 420
DB 361 VADSLVLTQDOLNSLAAYVLONRALDLTLTAERGTCFLGEECCYVYNOGSIYTEKYE 420
QY 421 IRDRIORAEELRNTGPMGLLSOMMPWILPFLGPIAAIITLLRPCLFNLVNVSSRI 480
DB 421 IRDRIORAEELRNTGPMGLLSOMMPWILPFLGPIAAIITLLRPCLFNLVNVSSRI 480
QY 481 EAVKLQMPKMQSKTKIYRRPLDRPASPRSDVNDIKGTPPEISAAQPLRPNSAGSS 538
DB 481 EAVKLQMPKMQSKTKIYRRPLDRPASPRSDVNDIKGTPPEISAAQPLRPNSAGSS 538

```

QY 241 SNTTNTNSOCIRMTWTPPTQIVCLPSGIFVCGTSAVRCJLNGSSSEMCFLSFVPMPTTY 300
|||||
Db 241 SNTTNTNSOCIRMTWTPPTQIVCLPSGIFVCGTSAVRCJLNGSSSEMCFLSFVPMPTTY 300
QY 301 TEODLYSYVYSKPRNKRVPILPFVIGAGYLGALGTGIGITTSSTQFYKLSOELNDGMR 360
|||||
Db 301 TEODLYSYVYSKPRNKRVPILPFVIGAGYLGALGTGIGITTSSTQFYKLSOELNDGMR 360
QY 361 VADSLVTLTODOLNSLAAYVLONRALDILTAERGCTCLFGECCYVVOGSIYTEKYE 420
|||||
Db 361 VADSLVTLTODOLNSLAAYVLONRALDILTAERGCTCLFGECCYVVOGSIYTEKYE 420
QY 421 IDRRIORRAEELRNTGPMGLLSQMPMWILPFLGPIAAIILLLFGPCIFNLVNFVSSRI 480
|||||
Db 421 IDRRIORRAEELRNTGPMGLLSQMPMWILPFLGPIAAIILLLFGPCIFNLVNFVSSRI 480
QY 481 EAVKLQMEPKMOSKTKIYRRPLDRPASPSVDNDIKGTPPEEISAAQPLLRPSAGSS 538
|||||
Db 481 EAVKLQMEPKMOSKTKIYRRPLDRPASPSVDNDIKGTPPEEISAAQPLLRPSAGSS 538

RESULT 4

US-09-719-554-26
; Sequence 26, Application US/09719554
; GENERAL INFORMATION:
; APPLICANT: ALLIET, Patrick
; APPLICANT: PERIN, Jean-Pierre
; APPLICANT: RIEGER, Francois
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCE AND DEDUCED PROTEIN SEQUENCE FAMILY WITH HU
; FILE REFERENCE: 200936050PCT
; CURRENT APPLICATION NUMBER: US/09/719,554
; CURRENT FILING DATE: 2001-12-26
; PRIOR APPLICATION NUMBER: PCT/FR99/01513
; PRIOR FILING DATE: 1999-06-25
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 26
; LENGTH: 540
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-719-554-26

Query Match 100.0%; Score 2871; DB 21; Length 540;
Best Local Similarity 100.0%; Pred. No. 9,2e-275;

Matches 538; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALPYHIFLFTVLLPSFTLTAPPCCRCMTSSSPYOEFILRMORPGNIDAPSYRSLSKGRP 60
|||||
Db 3 MALPYHIFLFTVLLPSFTLTAPPCCRCMTSSSPYOEFILRMORPGNIDAPSYRSLSKGRP 62
QY 61 TTTAHTHMRNCYHSATLTCMHANTHYWTGKMINPSCPGGLGYVVCWTYFTQGMSSGCGV 120
|||||
Db 63 TTTAHTHMRNCYHSATLTCMHANTHYWTGKMINPSCPGGLGYVVCWTYFTQGMSSGCGV 122
QY 121 QDOAREKHVKEYISOLTRVHGTSSPYKGLDLSKLEHETLRTHTRLVSLFNTTTLGLHEVSA 180
|||||
Db 123 QDOAREKHVKEYISOLTRVHGTSSPYKGLDLSKLEHETLRTHTRLVSLFNTTTLGLHEVSA 182
QY 181 QNPFTNCWICLPLNFRPYVSIYPVEOMNMFSTEINTTSVVLGVLVSNLEITHSNTLCVAF 240
|||||
Db 183 QNPFTNCWICLPLNFRPYVSIYPVEOMNMFSTEINTTSVVLGVLVSNLEITHSNTLCVAF 242
QY 241 SNTTNTNSOCIRMTWTPPTQIVCLPSGIFVCGTSAVRCJLNGSSSEMCFLSFVPMPTTY 300
|||||
Db 243 SNTTNTNSOCIRMTWTPPTQIVCLPSGIFVCGTSAVRCJLNGSSSEMCFLSFVPMPTTY 302
QY 301 TEODLYSYVYSKPRNKRVPILPFVIGAGYLGALGTGIGITTSSTQFYKLSOELNDGMR 360
|||||
Db 303 TEODLYSYVYSKPRNKRVPILPFVIGAGYLGALGTGIGITTSSTQFYKLSOELNDGMR 362
QY 361 VADSLVTLTODOLNSLAAYVLONRALDILTAERGCTCLFGECCYVVOGSIYTEKYE 420
|||||

Db 363 VADSLVTLTODOLNSLAAYVLONRALDILTAERGCTCLFGECCYVVOGSIYTEKYE 422
|||||
QY 421 IDRRIORRAEELRNTGPMGLLSQMPMWILPFLGPIAAIILLLFGPCIFNLVNFVSSRI 480
|||||
Db 423 IDRRIORRAEELRNTGPMGLLSQMPMWILPFLGPIAAIILLLFGPCIFNLVNFVSSRI 482
QY 481 EAVKLQMEPKMOSKTKIYRRPLDRPASPSVDNDIKGTPPEEISAAQPLLRPSAGSS 538
|||||
Db 483 EAVKLQMEPKMOSKTKIYRRPLDRPASPSVDNDIKGTPPEEISAAQPLLRPSAGSS 540
|||||

RESULT 5

US-09-949-016-11511
; Sequence 11511, Application US/09949016
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for windows version 4.0
; SEQ ID NO 11511
; LENGTH: 540
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-11511

Query Match 100.0%; Score 2871; DB 24; Length 540;
Best Local Similarity 100.0%; Pred. No. 9,2e-275;

Matches 538; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALPYHIFLFTVLLPSFTLTAPPCCRCMTSSSPYOEFILRMORPGNIDAPSYRSLSKGRP 60
|||||
Db 3 MALPYHIFLFTVLLPSFTLTAPPCCRCMTSSSPYOEFILRMORPGNIDAPSYRSLSKGRP 62
QY 61 TTTAHTHMRNCYHSATLTCMHANTHYWTGKMINPSCPGGLGYVVCWTYFTQGMSSGCGV 120
|||||
Db 63 TTTAHTHMRNCYHSATLTCMHANTHYWTGKMINPSCPGGLGYVVCWTYFTQGMSSGCGV 122
QY 121 QDOAREKHVKEYISOLTRVHGTSSPYKGLDLSKLEHETLRTHTRLVSLFNTTTLGLHEVSA 180
|||||
Db 123 QDOAREKHVKEYISOLTRVHGTSSPYKGLDLSKLEHETLRTHTRLVSLFNTTTLGLHEVSA 182
QY 181 QNPFTNCWICLPLNFRPYVSIYPVEOMNMFSTEINTTSVVLGVLVSNLEITHSNTLCVAF 240
|||||
Db 183 QNPFTNCWICLPLNFRPYVSIYPVEOMNMFSTEINTTSVVLGVLVSNLEITHSNTLCVAF 242
QY 241 SNTTNTNSOCIRMTWTPPTQIVCLPSGIFVCGTSAVRCJLNGSSSEMCFLSFVPMPTTY 300
|||||
Db 243 SNTTNTNSOCIRMTWTPPTQIVCLPSGIFVCGTSAVRCJLNGSSSEMCFLSFVPMPTTY 302
QY 301 TEODLYSYVYSKPRNKRVPILPFVIGAGYLGALGTGIGITTSSTQFYKLSOELNDGMR 360
|||||
Db 303 TEODLYSYVYSKPRNKRVPILPFVIGAGYLGALGTGIGITTSSTQFYKLSOELNDGMR 362
QY 421 IDRRIORRAEELRNTGPMGLLSQMPMWILPFLGPIAAIILLLFGPCIFNLVNFVSSRI 480
|||||
Db 423 IDRRIORRAEELRNTGPMGLLSQMPMWILPFLGPIAAIILLLFGPCIFNLVNFVSSRI 482
QY 481 EAVKLQMEPKMOSKTKIYRRPLDRPASPSVDNDIKGTPPEEISAAQPLLRPSAGSS 538
|||||

Db 483 EAVKLQEPKMKOSTKTYRRPLDRPASPRSDVNDIKGPPPEISAAQPLPNSAGSS 540

```
RESULT 6
US-09-719-554-35
; Sequence 35, Application US/09719554
; GENERAL INFORMATION:
; APPLICANT: ALBET, Patrick
; APPLICANT: PERIN, Jean-Pierre
; APPLICANT: RIEGER, Francois
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCE AND DEDUCED PROTEIN SEQUENCE FAMILY WITH HD
; FILE REFERENCE: 20036US0PCT
; CURRENT APPLICATION NUMBER: US/09/719,554
; PRIOR FILING DATE: 2001-12-26
; PRIOR APPLICATION NUMBER: PCT/FR99/01513
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 35
; LENGTH: 685
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (29)..(29)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (85)..(85)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (124)..(124)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (128)..(128)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (669)..(669)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (685)..(685)
; OTHER INFORMATION: Xaa is any amino acid
US-09-719-554-35
```

Query Match 100.0%; Score 2871; DB 21; Length 685;
Best Local Similarity 100.0%; Pred. No. 1.3e-274;
Matches 538; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MALPYHIFLFTVLLPSFTLTAPPCCRCMTSSSPYQDFLMRQRCNIDAPSYRSLSGTGP 60
Db 131 MALPYHIFLFTVLLPSFTLTAPPCCRCMTSSSPYQDFLMRQRCNIDAPSYRSLSGTGP 190
QY 61 TFTAHTMPRNCYSAATLCMHANTHYWTGKMINSPCGSLGVYTCWTFYQTGSDGCGV 120
Db 191 TFTAHTMPRNCYSAATLCMHANTHYWTGKMINSPCGSLGVYTCWTFYQTGSDGCGV 250
QY 121 ODAREKHVEVLSQLRVHGTSPPYKGLDLSKHETLRHTRLVSLFNTTLTGLHVEYA 180
Db 251 ODAREKHVEVLSQLRVHGTSPPYKGLDLSKHETLRHTRLVSLFNTTLTGLHVEYA 310
QY 181 QNTNCWICLPLNFRPVSLIPVEQWNNFSTEINTTSLVGLPLVSNLEITHTSNLTCKVF 240
Db 311 QNTNCWICLPLNFRPVSLIPVEQWNNFSTEINTTSLVGLPLVSNLEITHTSNLTCKVF 370
QY 241 SNTTNTNSOCIMWVTPPTQIVCLPSGIFVCGTSAYRCLNGSSEKCFSLVPPMTIY 300
Db 371 SNTTNTNSOCIMWVTPPTQIVCLPSGIFVCGTSAYRCLNGSSEKCFSLVPPMTIY 430
QY 301 TEODLYSYVSKRNRKRVPLLPVYIGAGVIGALGTGIGITTSYQFYFKLSQELNGMER 360
Db 431 TEODLYSYVSKRNRKRVPLLPVYIGAGVIGALGTGIGITTSYQFYFKLSQELNGMER 490
```

QY 361 VADSLVTLQDQNLNSLAAYVLLONRRALDLTAEERGTCFLGEBCCYVYNOGSIYTERKE 420
Db 491 VADSLVTLQDQNLNSLAAYVLLONRRALDLTAEERGTCFLGEBCCYVYNOGSIYTERKE 550
QY 421 IRRIOGRAEELRNTGPGWLLSQWMPWILPPLGLPAAITLLLLFGPCIFNLLVNFVSSRI 480
Db 551 IRRIOGRAEELRNTGPGWLLSQWMPWILPPLGLPAAITLLLLFGPCIFNLLVNFVSSRI 610
QY 481 EAVKLQEPKMKOSTKTYRRPLDRPASPRSDVNDIKGPPPEISAAQPLPNSAGSS 538
Db 611 EAVKLQEPKMKOSTKTYRRPLDRPASPRSDVNDIKGPPPEISAAQPLPNSAGSS 668

```
RESULT 7
PCT-US98-22034-4
; Sequence 4, Application PC/TUS9822034
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: Lavallee, Edward R.
; APPLICANT: Racle, Lisa A.
; APPLICANT: Werberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Evans, Cheryl
; APPLICANT: Spaulding, Vikki
; APPLICANT: Agostino, Michael J.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
; TITLE OF INVENTION: ENCODING THEM
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge
; STATE: MA
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US98/22034
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sprunger, Suzanne A.
; REGISTRATION NUMBER: 41,323
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8284
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 538 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US98-22034-4
```

Query Match 99.9%; Score 2868; DB 1; Length 538;
Best Local Similarity 99.8%; Pred. No. 1.8e-274;
Matches 537; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MALPYHIFLFTVLLPSFTLTAPPCCRCMTSSSPYQDFLMRQRCNIDAPSYRSLSGTGP 60
Db 1 MALPYHIFLFTVLLPSFTLTAPPCCRCMTSSSPYQDFLMRQRCNIDAPSYRSLSGTGP 60
QY 61 TFTAHTMPRNCYSAATLCMHANTHYWTGKMINSPCGSLGVYTCWTFYQTGSDGCGV 120
Db 61 TFTAHTMPRNCYSAATLCMHANTHYWTGKMINSPCGSLGVYTCWTFYQTGSDGCGV 120
QY 121 ODAREKHVEVLSQLRVHGTSPPYKGLDLSKHETLRHTRLVSLFNTTLTGLHVEYA 180
```

```
Db 121 ODOAREKHVKEVISQLTRHGTSSPYKGLDLSKLEHETLHTRLVSLFNTTLTGLHEVSA 180
OY 181 ONPTNCWICLPLNRPVYSIPVEQOMNNSFEINTNTSVLYGVLNSLETHTHNSLTCVKF 240
Db 181 ONPTNCWICLPLNRPVYSIPVEQOMNNSFEINTNTSVLYGVLNSLETHTHNSLTCVKF 240
OY 241 SNTTYTNSQCIRMTWTPPTQIVCLPSGIFVCGTSAYRCLNGSSSESMCFLSFLVPMPTIY 300
Db 241 SNTTYTNSQCIRMTWTPPTQIVCLPSGIFVCGTSAYRCLNGSSSESMCFLSFLVPMPTIY 300
OY 301 TEODLYSYIYSKPRNKRVPILPFIYAGVYLGALGTGIGITTSOTQYKLSOELNDMER 360
Db 301 TEODLYSYIYSKPRNKRVPILPFIYAGVYLGALGTGIGITTSOTQYKLSOELNDMER 360
OY 361 VADSLVTLTODOLNSLAAYVLONRALDILLTAERGTCCLFGECCCYVNOGSIYTEKVE 420
Db 361 VADSLVTLTODOLNSLAAYVLONRALDILLTAERGTCCLFGECCCYVNOGSIYTEKVE 420
OY 421 IRDRIORAEELRNTGPMGLSQMMPWILPFLGPLAAIILLFPGCIFNLVNFVSSRI 480
Db 421 IRDRIORAEELRNTGPMGLSQMMPWILPFLGPLAAIILLFPGCIFNLVNFVSSRI 480
OY 481 EAVKLOMEPKMOSKTKIYRRPLDRPASPSVDNDIKGTPEEISAAQPLLRPSAGSS 538
Db 481 EAVKLOMEPKMOSKTKIYRRPLDRPASPSVDNDIKGTPEEISAAQPLLRPSAGSS 538

RESULT 8
PCT-US98-24614-4
; Sequence 4, Application PC/TUS9824614
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: Lavallee, Edward R.
; APPLICANT: Collins-Racie, Lisa A.
; APPLICANT: Evans, Cheryl
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
; FILE REFERENCE: 6006B-PCT
; CURRENT APPLICATION NUMBER: PCT/US98/24614
; CURRENT FILING DATE: 1998-11-17
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 538
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US98-24614-4

Query Match          99.9%; Score 2868; DB 1; Length 538;
Best Local Similarity 99.8%; Pred. No. 1.8e-274;
Matches 537; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
Db 241 SNTTYTNSQCIRMTWTPPTQIVCLPSGIFVCGTSAYRCLNGSSSESMCFLSFLVPMPTIY 300
OY 301 TEODLYSYIYSKPRNKRVPILPFIYAGVYLGALGTGIGITTSOTQYKLSOELNDMER 360
Db 301 TEODLYSYIYSKPRNKRVPILPFIYAGVYLGALGTGIGITTSOTQYKLSOELNDMER 360
OY 361 VADSLVTLTODOLNSLAAYVLONRALDILLTAERGTCCLFGECCCYVNOGSIYTEKVE 420
Db 361 VADSLVTLTODOLNSLAAYVLONRALDILLTAERGTCCLFGECCCYVNOGSIYTEKVE 420
OY 421 IRDRIORAEELRNTGPMGLSQMMPWILPFLGPLAAIILLFPGCIFNLVNFVSSRI 480
Db 421 IRDRIORAEELRNTGPMGLSQMMPWILPFLGPLAAIILLFPGCIFNLVNFVSSRI 480
OY 481 EAVKLOMEPKMOSKTKIYRRPLDRPASPSVDNDIKGTPEEISAAQPLLRPSAGSS 538
Db 481 EAVKLOMEPKMOSKTKIYRRPLDRPASPSVDNDIKGTPEEISAAQPLLRPSAGSS 538
```

```
RESULT 9
PCT-US99-10915-4
; Sequence 4, Application PC/TUS9910915
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: Lavallee, Edward R.
; APPLICANT: Collins-Racie, Lisa A.
; APPLICANT: Evans, Cheryl
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: M. Sha
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
; FILE REFERENCE: 6006B.A1172A
; CURRENT APPLICATION NUMBER: PCT/US99/10915
; CURRENT FILING DATE: 1999-05-17
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 538
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US99-10915-4
```

```
Query Match          99.9%; Score 2868; DB 1; Length 538;
Best Local Similarity 99.8%; Pred. No. 1.8e-274;
Matches 537; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 MALPYHIFLFTVLLPSFTLTAPPCRCMTSSSPYOEFILRMORPGNIDAPSYRSLSKGRP 60
Db 1 MALPYHIFLFTVLLPSFTLTAPPCRCMTSSSPYOEFILRMORPGNIDAPSYRSLSKGRP 60
OY 61 TPTAHTHMPRNCYHSATLCMHANTHYWTGKMINPSCPGGLGYVTCWTFYPTQGMDSGGV 120
Db 61 TPTAHTHMPRNCYHSATLCMHANTHYWTGKMINPSCPGGLGYVTCWTFYPTQGMDSGGV 120
OY 121 ODOAREKHVKEVISQLTRHGTSSPYKGLDLSKLEHETLHTRLVSLFNTTLTGLHEVSA 180
Db 121 ODOAREKHVKEVISQLTRHGTSSPYKGLDLSKLEHETLHTRLVSLFNTTLTGLHEVSA 180
OY 181 ONPTNCWICLPLNRPVYSIPVEQOMNNSFEINTNTSVLYGVLNSLETHTHNSLTCVKF 240
Db 181 ONPTNCWICLPLNRPVYSIPVEQOMNNSFEINTNTSVLYGVLNSLETHTHNSLTCVKF 240
OY 241 SNTTYTNSQCIRMTWTPPTQIVCLPSGIFVCGTSAYRCLNGSSSESMCFLSFLVPMPTIY 300
Db 241 SNTTYTNSQCIRMTWTPPTQIVCLPSGIFVCGTSAYRCLNGSSSESMCFLSFLVPMPTIY 300
OY 301 TEODLYSYIYSKPRNKRVPILPFIYAGVYLGALGTGIGITTSOTQYKLSOELNDMER 360
Db 301 TEODLYSYIYSKPRNKRVPILPFIYAGVYLGALGTGIGITTSOTQYKLSOELNDMER 360
```

QY 361 VADSLVTLQDOLNSLAAYVLLQNRALDLTAERGTCFLGEBCCYVYNSGITYEYKE 420
DB 361 VADSLVTLQDOLNSLAAYVLLQNRALDLTAERGTCFLGEBCCYVYNSGITYEYKE 420
QY 421 IRDRIGRAEELRNTGPGWGLLSQWMPWLPFLGLPLAIIILLFPGCIFNLLVNFVSRI 480
DB 421 IRDRIGRAEELRNTGPGWGLLSQWMPWLPFLGLPLAIIILLFPGCIFNLLVNFVSRI 480
QY 481 EAVLQWEPKMQSKTKIYRRPLDRPASPRSDVNDIKGTPPEISAAQPLRPNSAGSS 538
DB 481 EAVLQWEPKMQSKTKIYRRPLDRPASPRSDVNDIKGTPPEISAAQPLRPNSAGSS 538

RESULT 10

US-08-976-110-4
Sequence 4, Application US/08976110
GENERAL INFORMATION:
APPLICANT: Jacobs, Kenneth
APPLICANT: McCoy, John M.
APPLICANT: Lavallee, Edward R.
APPLICANT: Racie, Lisa A.
APPLICANT: Merberg, David
APPLICANT: Treacy, Maurice
APPLICANT: Evans, Cheryl
APPLICANT: Spaulding, Vikki
APPLICANT: Agostino, Michael J.
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
TITLE OF INVENTION: ENCODING THEM
NUMBER OF SEQUENCES: 62
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: MA
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/976,110
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Sprunger, Suzanne A.
REGISTRATION NUMBER: 41,323
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8284
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 538 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-976-110-4

Query Match 99.9%; Score 2868; DB 13; Length 538;
Best Local Similarity 99.8%; Pred. No. 1.8e-274;
Matches 537; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALPYHFLFTVLLPSTLTAPPCCRMSSSPYQEFLLMORPGNIDAPSYNSLSKGP 60
DB 1 MALPYHFLFTVLLPSTLTAPPCCRMSSSPYQEFLLMORPGNIDAPSYNSLSKGP 60
QY 61 TFTAHTMPRNCYHSAFLCHNANTHWTKMINPSGCGGIGYVVCWTFYFOTGMSDGGV 120
DB 61 TFTAHTMPRNCYHSAFLCHNANTHWTKMINPSGCGGIGYVVCWTFYFOTGMSDGGV 120
QY 121 QDQAREKHVKEVISQLTRVHGTSSPYKGLDSLKHETLRTHRLVSLFNTTLNGLHEVSA 180

DB 121 QDQAREKHVKEVISQLTRVHGTSSPYKGLDSLKHETLRTHRLVSLFNTTLNGLHEVSA 180
QY 181 QNPTNCWICLPLNFRPYVSIPEEQMNFSTEINTSVLVQPLVSNLEITHTSNLTCKVF 240
DB 181 QNPTNCWICLPLNFRPYVSIPEEQMNFSTEINTSVLVQPLVSNLEITHTSNLTCKVF 240
QY 241 SNTTYTNSQCIRWTPPTQIVCLPSGIFPVCGTSAYNCLNGSSKSCFLSPVPPMTY 300
DB 241 SNTTYTNSQCIRWTPPTQIVCLPSGIFPVCGTSAYNCLNGSSKSCFLSPVPPMTY 300
QY 301 TEODLYSVIKSPRKRKRPVLPFVIAGVCLALGTGIGITTSQFYKLSOELNGDME 360
DB 301 TEODLYSVIKSPRKRKRPVLPFVIAGVCLALGTGIGITTSQFYKLSOELNGDME 360
QY 361 VADSLVTLQDOLNSLAAYVLLQNRALDLTAERGTCFLGEBCCYVYNSGITYEYKE 420
DB 361 VADSLVTLQDOLNSLAAYVLLQNRALDLTAERGTCFLGEBCCYVYNSGITYEYKE 420
QY 421 IRDRIGRAEELRNTGPGWGLLSQWMPWLPFLGLPLAIIILLFPGCIFNLLVNFVSRI 480
DB 421 IRDRIGRAEELRNTGPGWGLLSQWMPWLPFLGLPLAIIILLFPGCIFNLLVNFVSRI 480
QY 481 EAVLQWEPKMQSKTKIYRRPLDRPASPRSDVNDIKGTPPEISAAQPLRPNSAGSS 538
DB 481 EAVLQWEPKMQSKTKIYRRPLDRPASPRSDVNDIKGTPPEISAAQPLRPNSAGSS 538

RESULT 11

US-09-080-478-4
Sequence 4, Application US/09080478
GENERAL INFORMATION:
APPLICANT: Jacobs, Kenneth
APPLICANT: McCoy, John M.
APPLICANT: Lavallee, Edward R.
APPLICANT: Racie, Lisa A.
APPLICANT: Merberg, David
APPLICANT: Treacy, Maurice
APPLICANT: Evans, Cheryl
APPLICANT: Spaulding, Vikki
APPLICANT: Agostino, Michael J.
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
TITLE OF INVENTION: ENCODING THEM
NUMBER OF SEQUENCES: 62
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: MA
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/080,478
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/976,110
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Sprunger, Suzanne A.
REGISTRATION NUMBER: 41,323
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8284
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 538 amino acids
TYPE: amino acid

STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-080-478-4

Query Match 99.9%; Score 2868; DB 14; Length 538;
Best Local Similarity 99.8%; Pred. No. 1.8e-274;
Matches 537; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MALPYHIFLFTVLLPSFTLTAPPCCMTSSSPYQEFILRMORPGNIDAPSYRSLSKGRP 60
DB 1 MALPYHIFLFTVLLPSFTLTAPPCCMTSSSPYQEFILRMORPGNIDAPSYRSLSKGRP 60
QY 61 TTTAHTHMRNCYHSATLGMHANTHYWTGKMINPSCPGGLGYVVCWTYTTQGMSDGGCV 120
DB 61 TTTAHTHMRNCYHSATLGMHANTHYWTGKMINPSCPGGLGYVVCWTYTTQGMSDGGCV 120
QY 121 QOARERKHYKEYISQLTRVHGTSPPYKGLDLSKLTHTLTHRLVSLFNTTTLGLHEVSA 180
DB 121 QOARERKHYKEYISQLTRVHGTSPPYKGLDLSKLTHTLTHRLVSLFNTTTLGLHEVSA 180
QY 181 QNPNCWICLPLNFRPYVSIIPVEQWNNFSTEINTTSVLVGPLVSNLEITHNSNLTCVRF 240
DB 181 QNPNCWICLPLNFRPYVSIIPVEQWNNFSTEINTTSVLVGPLVSNLEITHNSNLTCVRF 240
QY 241 SMTTYTNSQCIKRWTPPTQIYCLPSGIFVCGTSAVRLNGSSSEMCFLSFLVPMPTIY 300
DB 241 SMTTYTNSQCIKRWTPPTQIYCLPSGIFVCGTSAVRLNGSSSEMCFLSFLVPMPTIY 300
QY 301 TEODLYSYISKPRNKRVPILPFIAGVYALGALGTIGITSTQFYKLSOELNDMER 360
DB 301 TEODLYSYISKPRNKRVPILPFIAGVYALGALGTIGITSTQFYKLSOELNDMER 360
QY 361 VADSLVTLDQNLNSLAAYVLONNRRLDILTAERGCTCLFGECCYVNOGSIYTERKYE 420
DB 361 VADSLVTLDQNLNSLAAYVLONNRRLDILTAERGCTCLFGECCYVNOGSIYTERKYE 420
QY 421 IRRDIORRAEELRNTGPMGLLSQMPWILPFLGPIAAITLLLFPGCIFNLVNFSSRI 480
DB 421 IRRDIORRAEELRNTGPMGLLSQMPWILPFLGPIAAITLLLFPGCIFNLVNFSSRI 480
QY 481 EAVKLOMERKMSKTIYRRPLDRPASPSDVNDIKGTPPEEISAAQPLLRPSAGSS 538
DB 481 EAVKLOMERKMSKTIYRRPLDRPASPSDVNDIKGTPPEEISAAQPLLRPSAGSS 538
```

RESULT 12

US-09-413-232-135
; Sequence 135, Application US/09413232
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: Lavalie, Edward R.
; APPLICANT: Collins-Racie, Lisa A.
; APPLICANT: Evans, Cheryl
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Bowman, Michael R.
; APPLICANT: Spaulding, Ylki
; APPLICANT: Catlin-Duckett, McKenough
; APPLICANT: Kelleher, Kerry S.
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
; FILE REFERENCE: GI 6000-10A
; CURRENT APPLICATION NUMBER: US/09/413,232
; CURRENT FILING DATE: 1999-10-06
; NUMBER OF SEQ ID NOS: 321
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 135
; LENGTH: 538
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-413-232-135

Query Match 99.9%; Score 2868; DB 18; Length 538;
Best Local Similarity 99.8%; Pred. No. 1.8e-274;
Matches 537; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MALPYHIFLFTVLLPSFTLTAPPCCMTSSSPYQEFILRMORPGNIDAPSYRSLSKGRP 60
DB 1 MALPYHIFLFTVLLPSFTLTAPPCCMTSSSPYQEFILRMORPGNIDAPSYRSLSKGRP 60
QY 61 TTTAHTHMRNCYHSATLGMHANTHYWTGKMINPSCPGGLGYVVCWTYTTQGMSDGGCV 120
DB 61 TTTAHTHMRNCYHSATLGMHANTHYWTGKMINPSCPGGLGYVVCWTYTTQGMSDGGCV 120
QY 121 QOARERKHYKEYISQLTRVHGTSPPYKGLDLSKLTHTLTHRLVSLFNTTTLGLHEVSA 180
DB 121 QOARERKHYKEYISQLTRVHGTSPPYKGLDLSKLTHTLTHRLVSLFNTTTLGLHEVSA 180
QY 181 QNPNCWICLPLNFRPYVSIIPVEQWNNFSTEINTTSVLVGPLVSNLEITHNSNLTCVRF 240
DB 181 QNPNCWICLPLNFRPYVSIIPVEQWNNFSTEINTTSVLVGPLVSNLEITHNSNLTCVRF 240
QY 241 SMTTYTNSQCIKRWTPPTQIYCLPSGIFVCGTSAVRLNGSSSEMCFLSFLVPMPTIY 300
DB 241 SMTTYTNSQCIKRWTPPTQIYCLPSGIFVCGTSAVRLNGSSSEMCFLSFLVPMPTIY 300
QY 301 TEODLYSYISKPRNKRVPILPFIAGVYALGALGTIGITSTQFYKLSOELNDMER 360
DB 301 TEODLYSYISKPRNKRVPILPFIAGVYALGALGTIGITSTQFYKLSOELNDMER 360
QY 361 VADSLVTLDQNLNSLAAYVLONNRRLDILTAERGCTCLFGECCYVNOGSIYTERKYE 420
DB 361 VADSLVTLDQNLNSLAAYVLONNRRLDILTAERGCTCLFGECCYVNOGSIYTERKYE 420
QY 421 IRRDIORRAEELRNTGPMGLLSQMPWILPFLGPIAAITLLLFPGCIFNLVNFSSRI 480
DB 421 IRRDIORRAEELRNTGPMGLLSQMPWILPFLGPIAAITLLLFPGCIFNLVNFSSRI 480
QY 481 EAVKLOMERKMSKTIYRRPLDRPASPSDVNDIKGTPPEEISAAQPLLRPSAGSS 538
DB 481 EAVKLOMERKMSKTIYRRPLDRPASPSDVNDIKGTPPEEISAAQPLLRPSAGSS 538
```

RESULT 13

US-09-791-537-31643
; Sequence 31643, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomedix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 31643
; LENGTH: 538
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-31643

Query Match 99.9%; Score 2868; DB 22; Length 538;
Best Local Similarity 99.8%; Pred. No. 1.8e-274;
Matches 537; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MALPYHIFLFTVLLPSFTLTAPPCCMTSSSPYQEFILRMORPGNIDAPSYRSLSKGRP 60
DB 1 MALPYHIFLFTVLLPSFTLTAPPCCMTSSSPYQEFILRMORPGNIDAPSYRSLSKGRP 60
QY 61 TTTAHTHMRNCYHSATLGMHANTHYWTGKMINPSCPGGLGYVVCWTYTTQGMSDGGCV 120
DB 61 TTTAHTHMRNCYHSATLGMHANTHYWTGKMINPSCPGGLGYVVCWTYTTQGMSDGGCV 120
```

QY 121 ODOAREKHVKEVISOQLRVHGTSSPYKGLDLSKIHETLRTHRLVSLFNTTLTGHEVSA 180
| | | | |
Db 121 ODOAREKHVKEVISOQLRVHGTSSPYKGLDLSKIHETLRTHRLVSLFNTTLTGHEVSA 180
QY 181 QNPNCWICLPLNRPVYSIPVPEQNMNFSTEINTSVLVGRLVSNLEITHTSNLTCVKF 240
| | | | |
Db 181 QNPNCWICLPLNRPVYSIPVPEQNMNFSTEINTSVLVGRLVSNLEITHTSNLTCVKF 240
QY 241 SNTYTTNSOCIRWVTPPTQIVCLPSGIFFCGTSARCLNGSSSESMCFLSFLVPPMTIY 300
| | | | |
Db 241 SNTYTTNSOCIRWVTPPTQIVCLPSGIFFCGTSARCLNGSSSESMCFLSFLVPPMTIY 300
QY 301 TEODLYSVISKPRNKRVPILPFVIGAGVLCALGTGIGITTSQFYFKLSQELNGDME 360
| | | | |
Db 301 TEODLYSVISKPRNKRVPILPFVIGAGVLCALGTGIGITTSQFYFKLSQELNGDME 360
QY 361 VADSLVTLQDQNLNSLAVALVQNRALDLTLAERGGTCLFLGEECCYYVNOGIVTEKYKE 420
| | | | |
Db 361 VADSLVTLQDQNLNSLAVALVQNRALDLTLAERGGTCLFLGEECCYYVNOGIVTEKYKE 420
QY 421 IRDRIORRAEELRMTGPGMLISQWMPWILPFLGLAAILLLLFPCIFNLVNFVSSRI 480
| | | | |
Db 421 IRDRIORRAEELRMTGPGMLISQWMPWILPFLGLAAILLLLFPCIFNLVNFVSSRI 480
QY 481 EAVKLQEPKMQSKTKIYRRPLDRPASPRSDVNDIKGTPPEISAQPLLRPNAGSS 538
| | | | |
Db 481 EAVKLQEPKMQSKTKIYRRPLDRPASPRSDVNDIKGTPPEISAQPLLRPNAGSS 538

RESULT 14
US-09-902-535-2
; Sequence 2, Application US/09902535
; GENERAL INFORMATION:
; APPLICANT: Kelch, Jr., James C.
; APPLICANT: McCoy, John M.
; APPLICANT: M. Sha
; TITLE OF INVENTION: Methods and compositions for diagnosing
; TITLE OF INVENTION: and treating preclampsia and gestational trophoblast
; FILE REFERENCE: GIN-6006B4
; CURRENT APPLICATION NUMBER: US/09/902,535
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: 60/216,657
; PRIOR FILING DATE: 2000-07-06
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 538
; TYPE: PR
; ORGANISM: Homo sapiens
US-09-902-535-2

Query Match 99.9%; Score 2868; DB 24; Length 538;
Best Local Similarity 99.8%; Pred. No. 1.8e-274;
Matches 537; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALPYHIFLFTVLLPSFTLAPPCRCMTSSPYOEFLMRQNRGNIDAPYSRLSKGTP 60
| | | | |
Db 1 MALPYHIFLFTVLLPSFTLAPPCRCMTSSPYOEFLMRQNRGNIDAPYSRLSKGTP 60
QY 61 TFTAHTHPRNCYSATLCHMANHYWTGKMINSPCGGLGVTWCMTYFTOTGMSDGGV 120
| | | | |
Db 61 TFTAHTHPRNCYSATLCHMANHYWTGKMINSPCGGLGVTWCMTYFTOTGMSDGGV 120
QY 121 ODOAREKHVKEVISOQLRVHGTSSPYKGLDLSKIHETLRTHRLVSLFNTTLTGHEVSA 180
| | | | |
Db 121 ODOAREKHVKEVISOQLRVHGTSSPYKGLDLSKIHETLRTHRLVSLFNTTLTGHEVSA 180
QY 181 QNPNCWICLPLNRPVYSIPVPEQNMNFSTEINTSVLVGRLVSNLEITHTSNLTCVKF 240
| | | | |
Db 181 QNPNCWICLPLNRPVYSIPVPEQNMNFSTEINTSVLVGRLVSNLEITHTSNLTCVKF 240

QY 241 SNTYTTNSOCIRWVTPPTQIVCLPSGIFFCGTSARCLNGSSSESMCFLSFLVPPMTIY 300
| | | | |
Db 241 SNTYTTNSOCIRWVTPPTQIVCLPSGIFFCGTSARCLNGSSSESMCFLSFLVPPMTIY 300
QY 301 TEODLYSVISKPRNKRVPILPFVIGAGVLCALGTGIGITTSQFYFKLSQELNGDME 360
| | | | |
Db 301 TEODLYSVISKPRNKRVPILPFVIGAGVLCALGTGIGITTSQFYFKLSQELNGDME 360
QY 361 VADSLVTLQDQNLNSLAVALVQNRALDLTLAERGGTCLFLGEECCYYVNOGIVTEKYKE 420
| | | | |
Db 361 VADSLVTLQDQNLNSLAVALVQNRALDLTLAERGGTCLFLGEECCYYVNOGIVTEKYKE 420
QY 421 IRDRIORRAEELRMTGPGMLISQWMPWILPFLGLAAILLLLFPCIFNLVNFVSSRI 480
| | | | |
Db 421 IRDRIORRAEELRMTGPGMLISQWMPWILPFLGLAAILLLLFPCIFNLVNFVSSRI 480
QY 481 EAVKLQEPKMQSKTKIYRRPLDRPASPRSDVNDIKGTPPEISAQPLLRPNAGSS 538
| | | | |
Db 481 EAVKLQEPKMQSKTKIYRRPLDRPASPRSDVNDIKGTPPEISAQPLLRPNAGSS 538

RESULT 15
US-10-016-249-4
; Sequence 4, Application US/10016249
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: Lavalley, Edward R.
; APPLICANT: Collins-Racie, Lisa A.
; APPLICANT: Evans, Cheryl
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: M. Sha
; TITLE OF INVENTION: Genetics Institute, Inc.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
; FILE REFERENCE: 6006B-A172A
; CURRENT APPLICATION NUMBER: US/10/016,249
; CURRENT FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: US/09/175,928
; PRIOR FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 538
; TYPE: PR
; ORGANISM: Homo sapiens
US-10-016-249-4

Query Match 99.9%; Score 2868; DB 26; Length 538;
Best Local Similarity 99.8%; Pred. No. 1.8e-274;
Matches 537; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALPYHIFLFTVLLPSFTLAPPCRCMTSSPYOEFLMRQNRGNIDAPYSRLSKGTP 60
| | | | |
Db 1 MALPYHIFLFTVLLPSFTLAPPCRCMTSSPYOEFLMRQNRGNIDAPYSRLSKGTP 60
QY 61 TFTAHTHPRNCYSATLCHMANHYWTGKMINSPCGGLGVTWCMTYFTOTGMSDGGV 120
| | | | |
Db 61 TFTAHTHPRNCYSATLCHMANHYWTGKMINSPCGGLGVTWCMTYFTOTGMSDGGV 120
QY 121 ODOAREKHVKEVISOQLRVHGTSSPYKGLDLSKIHETLRTHRLVSLFNTTLTGHEVSA 180
| | | | |
Db 121 ODOAREKHVKEVISOQLRVHGTSSPYKGLDLSKIHETLRTHRLVSLFNTTLTGHEVSA 180
QY 181 QNPNCWICLPLNRPVYSIPVPEQNMNFSTEINTSVLVGRLVSNLEITHTSNLTCVKF 240
| | | | |
Db 181 QNPNCWICLPLNRPVYSIPVPEQNMNFSTEINTSVLVGRLVSNLEITHTSNLTCVKF 240
QY 241 SNTYTTNSOCIRWVTPPTQIVCLPSGIFFCGTSARCLNGSSSESMCFLSFLVPPMTIY 300
| | | | |
Db 241 SNTYTTNSOCIRWVTPPTQIVCLPSGIFFCGTSARCLNGSSSESMCFLSFLVPPMTIY 300
QY 301 TEODLYSVISKPRNKRVPILPFVIGAGVLCALGTGIGITTSQFYFKLSQELNGDME 360
| | | | |

```
Db      |||:|||||
301 TEODLYNTVISKPRNKRVILPEFYLGAGLGTGIGITSTQFYKLSQELNGDMER 360
Qy      361 VADSLVTLODQUNSLAAVVLQNRRLDILTAERGCTCLFGECCYYVNOGSIYTEKYK 420
Db      361 VADSLVTLODQUNSLAAVVLQNRRLDILTAERGCTCLFGECCYYVNOGSIYTEKYK 420
Qy      421 IRDRIORRAEELRNTGPMGLLSQMPWILPFLGPIAAIILLLFGPCIFNLLVNFVSSRI 480
Db      421 IRDRIORRAEELRNTGPMGLLSQMPWILPFLGPIAAIILLLFGPCIFNLLVNFVSSRI 480
Qy      481 EAVKLQMEPKMOSKTKIYRRPLDRPASPRSDVNDIKGTPPEEISAAQPLLRPNASAGSS 538
Db      481 EAVKLQMEPKMOSKTKIYRRPLDRPASPRSDVNDIKGTPPEEISAAQPLLRPNASAGSS 538
```

Search completed: September 23, 2003, 15:05:30
Job time : 407 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model.

Run on: September 23, 2003, 14:55:22 : Search time 24 Seconds

(without alignments)
569.590 Million cell updates/sec

Title: US-10-069-883-1

Perfect score: 2871

Sequence: 1 MALPHYHFLFVTLPSFTLT.....PPEISANQPLLRPSAGSS 538

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 117788 seqs, 25409224 residues

Total number of hits satisfying chosen parameters: 117788

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/1/paa/PCF_NEW_COMB.pcp:*
2: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pcp:*
3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pcp:*
4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pcp:*
5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pcp:*
6: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pcp:*
7: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pcp:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	415	14.5	687	US-09-976-858-58	Sequence 58, App1
2	389.5	13.6	667	US-09-070-630-13	Sequence 13, App1
3	389	13.5	191	US-09-350-841A-1593	Sequence 1593, App1
4	101	3.5	860	PCF-US03-20322-217	Sequence 217, App
5	94.5	3.3	4641	US-60-490-890-2111	Sequence 2111, App
6	94	3.3	546	US-09-908-576-250	Sequence 250, App
7	92.5	3.2	592	PCF-US03-20322-217	Sequence 217, App
8	90.5	3.2	807	US-10-425-114A-43305	Sequence 43305, A
9	90.5	3.2	807	US-10-425-114A-45403	Sequence 45403, A
10	89.5	3.1	346	US-10-425-114A-38013	Sequence 38013, A
11	89	3.1	434	US-09-350-841A-1586	Sequence 1586, App
12	89	3.1	594	US-10-425-114A-5186	Sequence 5186, A
13	89	3.1	594	US-10-425-114A-62723	Sequence 62723, A
14	88.5	3.1	2538	PCF-US03-20082-49	Sequence 49, App1
15	88	3.1	459	US-60-487-610-2672	Sequence 2672, App
16	88	3.1	459	US-60-485-450-1720	Sequence 1720, App
17	88	3.1	721	US-10-425-114A-56506	Sequence 56506, A
18	87.5	3.0	820	US-60-495-114-1419	Sequence 1419, App
19	87.5	3.0	820	US-60-495-114-1420	Sequence 1420, App
20	87.5	3.0	820	US-60-495-114-1421	Sequence 1421, App
21	87.5	3.0	820	US-60-495-114-1422	Sequence 1422, App
22	87.5	3.0	833	US-10-326-824-2	Sequence 2, App1
23	87	3.0	415	US-10-425-114A-56903	Sequence 56903, A
24	87	3.0	453	US-60-487-610-2442	Sequence 2442, App
25	86.5	3.0	303	US-60-487-610-2389	Sequence 2389, App
26	86.5	3.0	303	US-60-495-135-416	Sequence 416, App

27	86.5	3.0	371	US-60-495-135-417	Sequence 417, App
28	86.5	3.0	883	PCF-US02-29560A-222	Sequence 222, App
29	86.5	3.0	2442	US-60-487-610-1629	Sequence 1629, App
30	86.5	3.0	2442	US-60-487-610-1748	Sequence 1748, App
31	85.5	3.0	528	PCF-US02-18638A-26	Sequence 26, App1
32	85.5	3.0	499	PCF-US03-26760-1358	Sequence 1358, App
33	85.5	3.0	651	US-10-296-734-8	Sequence 8, App1
34	84	2.9	480	PCF-US02-29560A-222	Sequence 2, App1
35	84	2.9	538	PCF-US03-26760-1358	Sequence 26, App1
36	84	2.9	669	US-60-490-890-1270	Sequence 1270, App
37	84	2.9	669	US-60-491-048-2	Sequence 2, App1
38	84	2.9	669	US-60-491-156-2	Sequence 2, App1
39	84	2.9	1165	US-09-837-961A-4	Sequence 4, App1
40	83.5	2.9	572	US-09-662-254B-22	Sequence 22, App1
41	83	2.9	350	US-10-425-114A-64542	Sequence 64542, A
42	82.5	2.9	246	US-10-654-102-57	Sequence 57, App1
43	82.5	2.9	370	US-10-296-115-1174	Sequence 1174, App
44	82	2.9	2004	US-60-490-890-1750	Sequence 1750, App
45	81.5	2.8	364	US-60-487-610-2673	Sequence 2673, App

ALIGNMENTS

RESULT 1
US-09-976-858-58
Sequence 58, Application US/09976858
GENERAL INFORMATION:
APPLICANT: Gish, Kurt C.
APPLICANT: Mack, David H.
APPLICANT: Wilson, Keith E.
APPLICANT: Afari, Daniel
APPLICANT: Peter, Hevezi
TITLE OF INVENTION: Methods of Diagnosis of Prostate Cancer, Compositions and Methods of Screening for Modulators of Prostate Cancer
FILE REFERENCE: 05882.0183.NPUS00
CURRENT APPLICATION NUMBER: US/09/976,858
CURRENT FILING DATE: 2001-10-12
PRIOR APPLICATION NUMBER: 60/276,791
PRIOR FILING DATE: 2001-03-16
PRIOR APPLICATION NUMBER: 60/288,589
PRIOR FILING DATE: 2001-05-04
PRIOR APPLICATION NUMBER: 60/276,888
PRIOR FILING DATE: 2001-03-16
PRIOR APPLICATION NUMBER: 60/286,214
PRIOR FILING DATE: 2001-04-24
PRIOR APPLICATION NUMBER: 60/281,922
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: 60/263,957
PRIOR FILING DATE: 2001-01-24
NUMBER OF SEQ ID NOS: 294
SOFTWARE: PatentIn version 3.2
SEQ ID NO 58
LENGTH: 687
TYPE: PRT
ORGANISM: human organism
US-09-976-858-58
Query Match 14.5% Score 415; DB 5; Length 687;
Best Local Similarity 26.7% Pred. No. 7.2e-30;
Matches 124; Conservative 76; Mismatches 182; Indels 82; Gaps 20;
DB 70 RNCVHATLTCM-----ANTHY-----WTCCKMINPSCPGSLGV-----VCMYFTONG-MS 115
245 KNC-INPLTTVHPNNAOMTYGMSWLRLYIFEDVGMFTTOKITLVGMSKPPIGPLT 303
116 DGGVQDQAREKHEVKEVISQLTFVHGTSSPYKGLDLSKLTETLRTLRVLSLENTLTL 175
304 DLG---DPIFOKRPDKV---DLTVPLPLVPRQLQOOLHPLSLM-----STLGSV 348
176 HEV-----SAONPTNCICLPLNFRPVSPVPEQWNN--FSTEINTSVLGVPLVSNLEI 229
349 HHLNLTQPKLQDDCWICLKAKPPYVVGVEATLTKRGPLSCHTRPRALITGDV----- 402

370 DQNSLA VV LQNRALDLTAERGTCLEFGECCYYNOSGI VTEKVK EIRDRIOBRA 429

0Y B6 YWTGKMINSKPGIGVTVCWTTYETTOTGMSDGGGVODABEKHKYKEVTSQITRVHGTSSP 145

Db 333 -----PROACRF-----GGRREAMEOE--VKOTLVQHPR----- 360
QY 146 YKGL-DLSKHEH-----LRTHRLVSLFNTLTGLGHEVSAQNPTN---CWIC 189
Db 361 YGINDTGKINTKPCAGSDPEYAFMTWCRGFEFLCANMTWFLNWEEDNKNQTRRNCHIK 420
QY 190 LPLNRPVY--SIPPEQNNFSTEINTSVLGLVLSNLEI-----THTSNLTCVKSMT 243
Db 421 QIINTMHKVKKNVLPREGELACESTVSI-----IANIDIKNTH-----NI 466
QY 244 TTTTNSQCIRMTPTPTQIYCLPSGIFVCGTSATYRCLNGSSSESMCFSLFLVPMPTTTEQ 303
Db 467 TESAEEV-----AELYRIELGDYK-----LIETPTIGFAPTDQ 498
QY 304 DLSYVYSKP-RNRKRPILPFYIG-AGVIGALGTGIGITTSQFYFKLSQELNGSMERY 361
Db 499 RRTS---STPVRNKRK---VFVLGFLGFLATAGSANG----- 529
QY 362 ADSLVTLQDOLNSLAAVLQNRALDLTAER-----GCT----- 396
Db 530 AKSLTLSAQSRLLAGIYQOQQLDLYVRQOEMLLTVWGTFKNLQARTALEKYLKHQA 589
QY 397 -----CLFLGEECCY-----YVNOQ-----GIYTKYKEIRDR 425
Db 590 QUNSMQCAF--ROYCHTTPVWVNDLSLSPDKMNTWQEMEKQVYLEANISOSLEFAQIOQ 647
QY 426 ORRAEELRNTGPMGLISQMPMLLPFLGPLAIIILLFGPCIFNLVNFVSRIEAVL 485
Db 648 EKMYELQKLNKSDILGNMF-----DLTSVWKIYQY---VHIVGIIALRIAYV 696
QY 486 QMEPKMSKTKIYRRPLDRPAS--PRSDYNDIKGTPPEE 522
Db 697 QLLSRFR---KGYRPVSSPPGLQOIHIHKRGQANE 732

RESULT 5
US-60-490-890-2111
; Sequence 2111, Application US/60490890
; GENERAL INFORMATION:
; APPLICANT: Li, Martha
; APPLICANT: Rupnow, Brent A.
; APPLICANT: Webster, Kevin R.
; APPLICANT: Jackson, Donald
; APPLICANT: Wong, Tai W.
; TITLE OF INVENTION: BIOMARKERS OF CYCLIN-DEPENDENT KINASE MODULATION
; FILE REFERENCE: D0310 BSP
; CURRENT APPLICATION NUMBER: US/60/490,890
; NUMBER OF SEQ ID NOS: 2779
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 2111
; LENGTH: 4641
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-490-890-2111

Query Match 3.3%; Score 94.5; DB 7; Length 4641;
Best Local Similarity 19.0%; Pred. No. 26;
Matches 111; Conservative 64; Mismatches 165; Indels 243; Gaps 25;

QY 70 RNCYHSATLCMANHTHYMTGKMIINPSCPGICAGTVCMTFTQTGMDGGGVODQ-AREKH 128
Db 1502 QECNHTFTACPHS---FYPTPALQWMC-----LQDLINCLDQDQIOEAN 1541
QY 129 VEEVISQLTRVHGTSPPYGLDISKIHETL-RTHTRLVSLFNTTLG-----LHEVSA 180
Db 1542 FKTSSSRL-----LAAVMSALCHTSVKLTSTFPIAYDGEVLLRSIRKOVST 1587
QY 181 ONPTNCWICLPL-----NRRP-----YVSLPPEQNNN-----TS 210
Db 1588 ENDSTLVHRPPLLVAMHEKLSOSEENISGNTSREVLKMLVIVLVRNLSRRENEELS 1647
QY 211 TE-INTSVLVGLVSNL-----EITHTSNLTCVKSMTYTTNSQCIAMWTPTPTQIY 262

Db 1648 SHLVSNFTCGLLASIVSELASALGSEVDGILNSLHSVKASANRPTKTSQGRSMTG----- 1702
QY 263 CLPSGIFVCGTSATYRCLNGSSSESKCF-----LSFLVPPMT 298
Db 1703 -----NOSPDAICFSYDKPGIYVVGSGSVYGGGIIHEYELEVLVDSE 1744
QY 299 -----IYTEQDLYSVVISKPRKRPILPFYIGAGVIGALGT----- 335
Db 1745 HAGDSTSHRMTSLBYKGTYYTDDSPSDIAETRLDKVYVLEKNVYAVRLRNYGSRGAN 1804
QY 336 GIGGITY-----ST-----QFYKLSQELNGDMERYADSLYT 367
Db 1805 GGGGMYTTCPCPGVTFPTFCSLNSNGTNOTRQIPIQLIYR--SEFDD----- 1852
QY 368 LQDOLNSLAAVLQNR-RLADLTLAERGTCFLGEECCYVNOQSIYTKYKEIRDR 426
Db 1853 LOSQLLSKANEEDKNCSTRAL-----SVSGTVVVAASKDL 1887
QY 427 RR-----AEELRNTGPMGLISQMPMLLPFLGPLAIIILLFGPCIFNLVNFVSRR 479
Db 1888 RALAVDADDIPELLSS--SLFSMLLPILITAYIGVAAAI-----PKVAVE 1931
QY 480 IEAVKLOMEPKMSKTKIYRRPLDRPASPRSDYNDIKGTPPEE 522
Db 1932 VEGLVQQLLPVAILNQKYAP--AFNPQSTSTDTGNQPEQ 1971

RESULT 6
US-09-908-576-250
; Sequence 250, Application US/09908576
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Bolstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gertsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/908,576
; CURRENT FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US/09/665,350B
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594

```

: PRIOR FILING DATE: 1999-09-08
: PRIOR APPLICATION NUMBER: PCT/US99/20944
: PRIOR FILING DATE: 1999-09-13
: PRIOR APPLICATION NUMBER: PCT/US99/21090
: PRIOR FILING DATE: 1999-09-15
: PRIOR APPLICATION NUMBER: PCT/US99/21547
: PRIOR FILING DATE: 1999-09-15
: PRIOR APPLICATION NUMBER: PCT/US99/23089
: PRIOR FILING DATE: 1999-10-05
: Remaining Prior Application data removed - See file Wrapper or PALM.
: NUMBER OF SEQ ID NOS: 423
: SEQ ID NO: 250
: LENGTH: 546
: TYPE: PRT
: ORGANISM: Homo Sapien
: US-09-908-576-250

```

```

Query Match 3.3%; Score 94; DB 5; Length 546;
Best Local Similarity 20.3%; Pred. No. 1.5;
Matches 101; Conservative 67; Mismatches 145; Indels 184; Gaps 28;

```

```

QY 3 LPYHIFLFT-----VLPSEFTLAPPCRCMTSSSPYGFELMRMQRG 45
DB 164 LELHFLMSGIPDYFDLVELEVLKLEIPDYTI--PPSIQOLGKKE---LMLYHTAA 217
QY 46 NIDAPSYRSLSKGPT---FTANTHMRNCYHSATLCMHANTHYMTGKMINPSCPGIG 101
DB 218 KIEAPALAFLEHNLRAHIFKFTDIKEIPLWISLKT---ELHLTG--NLSENNRY 270
QY 102 VTVCMTYFTOTGMDGGGVQOAREK--HKEVISQLTRV-----HGT-- 142
DB 271 IVI-----DLRELRLKYLRLKLSNLSKLPQVYTDYVHLQKLSINNEGKTL 317
QY 143 ---SSPYKGLDLSKLEHTRHTRLV--SLFNTLTGHEVSAONPTNCWICLPNRPY 197
DB 318 IVLNSLKKMNLTEL--ELIRCDLERIPRISIT--SLHNOELDKD----- 359
QY 198 VSIPEQWNNFSTEINTSYLVGPELVSNLEITHSNLTCVK--FSNTTY-----TTN 248
DB 360 -----NNLKT-----IEEISFOHLHRLTCLKLMWNHAIYPIQIGNLTN 399
QY 249 SOCIARWYPTQIYCLPSGIFVCGTSYRCLNSSSEKCLSLVPPMTYITQDLYSY 308
DB 400 LE--RLYLNKRIKIKITQIFY--CRKLY--LDLSHNNLFTL---PADIGLQNLONT 449
QY 309 VISKPRNKRPV-----ILPVIAGVGLGALGTGIGITSTQFYKLSOELGDMER 360
DB 450 AITANRIETLPELFCQCKRLALHGNVLOSLSRVELNLNLIQI-----ELRG----- 499
QY 361 VADSLVTLQDQNSLAVALYQNRALDLTPAERGGTCLFLGEECCYVYNOSGIYTER-- 417
DB 500 -----NR--LLECLPVE-----LGE--CPILKSGLVVEEDLF 527
QY 418 ---VKEIRRIORRAEE 431
DB 528 NTLPEVKEKRLMRADKE 544

```

```

RESULT 7
PCT-IL00-00358-3
: Sequence 3, Application PC/TIL0000358
: GENERAL INFORMATION:
: APPLICANT: pecker, lrls
: TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ENCODED THEREBY DISTANTLY HOMOLG
: FILE REFERENCE: 00/20105
: CURRENT APPLICATION NUMBER: PCT/IL00/00358
: CURRENT FILING DATE: 2002-06-19
: PRIOR APPLICATION NUMBER: 60/140,801
: PRIOR FILING DATE: 1999-06-25
: NUMBER OF SEQ ID NOS: 25
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 3

```

```

: LENGTH: 592
: TYPE: PRT
: ORGANISM: Homo sapiens
: PCT-IL00-00358-3

```

```

Query Match 3.2%; Score 92.5; DB 1; Length 592;
Best Local Similarity 20.7%; Pred. No. 2.4;
Matches 62; Conservative 41; Mismatches 87; Indels 109; Gaps 15;

```

```

QY 121 QOQAREKIVKEVISLTVHGTSSPYKGLDLSKLEHTRHTRLVSLFN--TTLGLHAYS 179
DB 178 REKAQOMHLVLKKEOF-----SMTYSMLIL-----TARSIDKLNFADCGSLHIF 223
QY 180 A-----QNPNCWICLPNRPYISIPVEQWNNFSTEINTSYLVGPELVSNLEITHRSN 234
DB 224 ALNLRKRPNNNSW-----NNSSA 241
QY 235 LTCVKFSNTYTTNSQICRWYPTQIYCLPSGIFVCGTSAYRCLNCS--SESNCPLSF 292
DB 242 LSLKYS---ASKKYNISW-----ELGNEPNNYRTMG---BAVNSQGLKDYIQLKS 288
QY 293 LVPPMTYITBEDLYSYVSKPRNKRPVILPPIVIGAGVGLGALGTGIGITSTQFYKLSQ 352
DB 289 LLOPRTIYRSASLYGPNIGRPRKNVIALD---GEKRVAGSVDAVT--WQHCY---- 337
QY 353 ELNGDMERVADSLV-----TLQDQNSLAVALY-----LQNRALDLTPAERGGT 396
DB 338 -IDGRVYKVMDFLTKRLDLTISDQIRKIQKVVNTYTPCKKIMLG-----VYITSAGT 390

```

```

RESULT 8
US-10-425-114A-43305
: Sequence 43305, Application US/10425114A
: GENERAL INFORMATION:
: APPLICANT: Liu, Jindong
: APPLICANT: Zhou, Yihua
: APPLICANT: Kovacic, David K.
: APPLICANT: Screen, Steven E
: APPLICANT: Tabaska, Jack E
: APPLICANT: Cao, Yongwei
: TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
: FILE REFERENCE: 38-21(5313)B
: CURRENT APPLICATION NUMBER: US/10/425,114A
: CURRENT FILING DATE: 2003-04-28
: NUMBER OF SEQ ID NOS: 73128
: SEQ ID NO 43305
: LENGTH: 807
: TYPE: PRT
: ORGANISM: Zea mays
: FEATURE:
: OTHER INFORMATION: Clone ID: 700096110_FLI_pap
: US-10-425-114A-43305

```

```

Query Match 3.2%; Score 90.5; DB 6; Length 807;
Best Local Similarity 21.6%; Pred. No. 5.5;
Matches 77; Conservative 38; Mismatches 113; Indels 129; Gaps 20;

```

```

QY 54 SLKSGTPPTPAH--THMPRNCY-----HSATLCMHANTH---YTGKMINPSCGGL 100
DB 263 AIPSGTETOPYNDTOMADNPRYSCSLDDOSSGCTGTONSVHEPILYMTGAV-----E 314
QY 101 GVTVCMWYPTQIGMSDGGVQOQAREKIVKEVISLTVHGTSSP--YKGLDLSKLEHTR 158
DB 315 GSAVSYPQTLPG-----ALQSGLMRPKCKLSRKYVPTEQNALSOHL 357
QY 159 RTHTRLVSLFN---TTLTGHEVSAONPTN-----CHICLP-- 191
DB 358 RGMRRR--SLFNEKAGISNKGVDKRSDHHPVWSTPKCKTISGDNKSKPLRTPPC--ALPGI 414
QY 192 -LNFRRPVVIP---VEQWNNFSTEINTSYLVGPELVSNLEITHRSNLT 236
DB 415 GLHNLALATIPKEKIVRE---IOSTINESNLIGPAGSSPAPEQNIINDPOTYTDVA 471

```


THIS PAGE BLANK (USPTO)

us-10-069-883-1.ral

Tue Sep 23 15:54:18 2003

Db 122 WTSFHFGLPYVMPILGLPLCLLVLSFGPIFNKMTFIKHQIESIQ 168

RESULT 15
US-08-484-223B-107 Application US/08484223B

Sequence 107, Patent No. 6020459
GENERAL INFORMATION:
APPLICANT: Bolognesi, Dani P.
APPLICANT: Matthews, Thomas J.
APPLICANT: Wild, Carl T.
APPLICANT: Barney, Shawn O.
APPLICANT: Lambert, Dennis M.
APPLICANT: Peteway, Stephen R.
APPLICANT: Langlois, Alphonse J.
TITLE OF INVENTION: MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
TITLE OF INVENTION: TRANSMISSION
TITLE OF INVENTION: TRANSMISSION
NUMBER OF SEQUENCES: 245
CORRESPONDENCE ADDRESS: LLP
ADDRESS: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,223B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Cornuzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7872-029
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 107:
SEQUENCE CHARACTERISTICS:
LENGTH: 192 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein

US-08-484-223B-107 Score 412; DB 3; Length 192;

Query Match Similarity 44.3%; Pred. No. 6e-34; 51; Indels 0; Gaps 0;
Best Local Similarity 42; Mismatches 51; Indels 0; Gaps 0;

Matches 74; Conservative 42; Mismatches 51; Indels 0; Gaps 0;
318 VPILPEVIGAGVLCALGTGIGITTSQFYKLSQELNGMERVADSLVTLDDQINSIAA 377
2 IOLPLFVGLGITTAIVSTGAAGAGVITQYTKLSHQSLSDVQALISSTIODLDQVDSLAE 61
378 VVIONRAIDLTTARGGTCLFLGEGCYVYVNOGSIYKVEIKRDRIORARELRNTP 121
62 VVIONRGLDLTAEAGGICLAEKCCFYANKSGIVRDKIKNLDDLRERRROLIDNPF 121
438 WLLSQMFWLIPFELGPLAAIILLFPCFENLVFVSSRIEAVK 484
122 WTSFHFGLPYVMPILGLPLCLLVLSFGPIFNKMTFIKHQIESIQ 168

Search completed: September 23, 2003, 14:58:38
Job time 33 secs

Tue Sep 23 15:54:18 2003

us-10-069-883-1.ra1

Page 7

OY 270 FVCGTATYRCLNGS-----SSSMCLSLVLPMTITREODLYSVISNPKRKRYPI---LP 322
DB 1065 MACSTGLTFCISTITLLMTDYCVLELMPRTVTHSPSYVGLFERSNHRKREPSTLTA 1124
OY 323 FVIGAGVYALGATGIGITTT---STOFYKLSOBLNDMDERVADSLVTLDDOQNSLAAV 379
DB 1125 LLLGLTGMGIAAGICTGTTLAMTAOFOOLAAYODDLREKXSISNLSLSLSEVY 1184
OY 380 LONRRALDLTLTAREGTCGLFGECCYVYVOSGIVTEKYEKRIORRAEELRNTGTP 437
DB 1185 LONRRGLDLTLAEGGICLALOEKCCFYANKSGIVRDKIKNLODERRROLIDNP 121
OY 439 GLLSOMPMPLIFLPGPLAIIILLFGPCIFNLVNFVSSRIEAVK 484
DB 1245 EGLFNR-SPMFTTLLSTINGPLVLLMLLFGPCILNRYOVFKRISVYQ 1294

RESULT 13
US-08-486-099-107
Sequence 107, Application US/08486099
Patent No. 6013263
GENERAL INFORMATION:
APPLICANT: Bolognesi, Daniel P.
APPLICANT: Matthews, Thomas J.
APPLICANT: Wild, Carl T.
APPLICANT: Barney, Shawn O.
APPLICANT: Lambert, Dennis M.
APPLICANT: Peteway, Stephen R.
TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF
TITLE OF INVENTION: MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HEPATITIS
NUMBER OF SEQUENCES: 209
CORRESPONDENCE ADDRESS:
ADDRESS: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/486.099
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7872-031
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
INFORMATION FOR SEQ ID NO: 107:
SEQUENCE CHARACTERISTICS:
LENGTH: 192 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-486-099-107

Query Match
Best Local Similarity 14.48; Score 412; DB 3; Length 192;
Matches 74; Conservative 42; Mismatches 51; Indels 0; Gaps 0;
OY 318 VPLIPVIGAGVYALGATGIGITTTSTOFYKLSOBLNDMDERVADSLVTLDDOQNSLAA 377

DB 2 IOLIPLVGIGITTAVSTGAAGVSIQYTKLSHOLISDVQAISSTIODDOVDLSLAE 61
OY 378 VYONRRALDLTLTAREGTCGLFGECCYVYVOSGIVTEKYEKRIORRAEELRNTGTP 437
DB 62 VYONRRGLDLTLAEGGICLALOEKCCFYANKSGIVRDKIKNLODERRROLIDNP 121
OY 438 WGLSOMPMPLIFLPGPLAIIILLFGPCIFNLVNFVSSRIEAVK 484
DB 122 WTSRHFLLFVYVPLGILLCLLVLSRGPILFNKIMLFTIKKIQSLIO 168

RESULT 14
US-08-360-107A-117
Sequence 117, Application US/08360107A
Patent No. 6012536
GENERAL INFORMATION:
APPLICANT: Bolognesi, Daniel P.
APPLICANT: Matthews, Thomas J.
APPLICANT: Wild, Carl T.
APPLICANT: Barney, Shawn O.
APPLICANT: Lambert, Dennis M.
APPLICANT: Peteway, Stephen R.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION
TITLE OF INVENTION: OF MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
NUMBER OF SEQUENCES: 149
CORRESPONDENCE ADDRESS:
ADDRESS: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/360.107A
FILING DATE: 20-DEC-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7872-013
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
INFORMATION FOR SEQ ID NO: 117:
SEQUENCE CHARACTERISTICS:
LENGTH: 192 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-360-107A-117

Query Match
Best Local Similarity 14.48; Score 412; DB 3; Length 192;
Matches 74; Conservative 42; Mismatches 51; Indels 0; Gaps 0;
OY 318 VPLIPVIGAGVYALGATGIGITTTSTOFYKLSOBLNDMDERVADSLVTLDDOQNSLAA 377

DB 2 IOLIPLVGIGITTAVSTGAAGVSIQYTKLSHOLISDVQAISSTIODDOVDLSLAE 61
OY 378 VYONRRALDLTLTAREGTCGLFGECCYVYVOSGIVTEKYEKRIORRAEELRNTGTP 437
DB 62 VYONRRGLDLTLAEGGICLALOEKCCFYANKSGIVRDKIKNLODERRROLIDNP 121
OY 438 WGLSOMPMPLIFLPGPLAIIILLFGPCIFNLVNFVSSRIEAVK 484

```

QY      32  FVIGAAVLALAGTGTGIGIT---SPOFYKKLSOELNGMEXVADSLVTLDDQNSLAAY 379
Db      445  LILGGJTMGGIAGIGTGTTLMAAQOFOOLAAVQDDLRIVERYSISLTKSLTSEVV 504
QY      380  LONRRALDILLTARGGTCTLGECCCYVYNOSGIVTEKVERIRRIQRAEELRNTGPW- 438
Db      505  LONRRALDILLTKEGGLCAALKEECCFYADHGTGIVRSMALRLRLNQRKLFESTGWF 564
QY      439  -GLLSOMMP-----ILPFLGPLLAITLLTLSPCTFNLNVFVSRIEAYK 484
Db      565  EGEFNR-SWPFETTLSTINGPLIVLIMLLTGPCPLNRLLVQFVADRISVQ 614

```

```

RESULT 10
US-09-315-127-3
: Sequence 3, Application US/09315127
: Patent No. 6448390
: GENERAL INFORMATION:
: APPLICANT: The University of Tennessee, c/o Richard Cox
: TITLE OF INVENTION: Stable Envelope Proteins for Retroviral, Viral and
: FILE OF INVENTION: Liposome Vectors and Use in Gene and Drug Therapy
: PILE REFERENCE: 44137-5023, U. of Tennessee
: CURRENT APPLICATION NUMBER: US/09/315,127
: NUMBER OF SEQ ID NOS: 23
: SOFTWARE: Patentln Ver. 2.0
: SEQ ID NO 3
: LENGTH: 632
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence: SEQ ID NO.2,
: OTHER INFORMATION: envelope protein produced by retroviral vector of
: US-09-315-127-3

```

[illegible]

RESULT 11
US-09-309-572-14
Sequence 14, Application US/09309572
Patent No. 6440730
GENERAL INFORMATION:
APPLICANT: Helmholtz-Pette-Institut
TITLE OF INVENTION: Retroviral hybrid vectors
FILE REFERENCE: P50489
pseudotyped with ICNV

```

:
: CURRENT APPLICATION NUMBER: US/09/309,572
:
: CURRENT FILING DATE: 1999-05-11
:
: EARLIER APPLICATION NUMBER: DE 198 56 463
:
: EARLIER FILING DATE: 1998-11-26
:
: NUMBER OF SEQ ID NOS: 24
:
: SOFTWARE: PatentIn Ver. 2.0
:
: SEQ ID NO 14
:
: LENGTH: 665
:
: TYPE: PRT
:
: ORGANISM: Moloney murine leukemia virus
:
: FEATURE:
:
: OTHER INFORMATION: env protein
:
: US-09-309-572-14

```

	Query Match	14.6%	Score 418.5:	DB 4;	Length 665;
	Best Local Similarity	30.2%;	Pred No. 1e-33;		
	Matches	106;	Conservative	63;	Mismatches 143; Indels 39; Gaps 9
QY	160	THTRLVSLFNTTLTGLHEVSANONPTNCWICPLNRPVPVSPVEOMNNFTSEINTTSVL	219		
Db	310	TENRLLNVDCAYQALNLTSPDKNGEHCILCAGPPRYEGVALGCTSYNHSAPANCVA	369		
QY	220	VGPLVSNLEIT-----HTSNLVCVKFSNNITYTTNSOCIRKWTPTQTIVCLPSCIF	269		
Db	370	SQHLLLTSEVTGGGICIGAVPKTHOALC-----NTQTSSR-----GSYLVAPTGTM	417		
QY	270	FVCGMSAARCLNGS-----SESMCFLSLFPVPMITITBEDLYSVYSKRPNRKRVPL---LP	322		
Db	418	WAGSTGLTPCISTTYLTNLNTDYCYLVELMPRTYHSPSYGVLFERSNRHKREPVSTLTA	477		
QY	323	FVISAGVLGALTGTIGIGITT---STOFKYKLSOEIENGMERADSLVTLQDOLNSIAVV	379		
Db	478	LLOGGITMGGAIGAGTGTCITLTMANQQOQLAAVODDLREVEXISLNLSLSLSEVY	537		
QY	380	LQNRRPALDLTAERGCITCLFGEECCYVYNOSGIYTERKEIKRIDRIQRABEALRNTGPV-	438		
Db	538	LQNRGRDILLFLFKEGGICALMKKECCFYADHTGIVRDGMARKRELNRORUKLFESTQWF	597		
QY	439	-GLISSQMPPV-----ILPRFLGPLAAITLLLLRGPCFLNLVNPFVSSRIEAKV	484		
Db	598	EGLERN-SPEMTTLITSTMGPLVIALMLILFCGPCILNLRVDFVKRISRIVQ	647		

RESULT 12
 US-09-554-572-26
 : Sequence 26: Application US/09554572
 : Patent No. 6573891
 : GENERAL INFORMATION:
 : APPLICANT: NATURE TECHNOLOGY, INC.
 : TITLE OF INVENTION: CHIMERIC VIRAL PACKAGING SIGNAL WITHOUT GAG GENES
 : TITLE OF INVENTION: SEQUENCES
 : FILE REFERENCE: 228.00030201
 : CURRENT APPLICATION NUMBER: US/09/554,572
 : CURRENT FILING DATE: 2000-09-18
 : NUMBER OF SEQ. ID NOS: 26
 : SOFTWARE: PatentIn Ver. 2.0
 : SEQ. ID NO: 26
 : LENGTH: 1312
 : TYPE: PRT
 : ORGANISM: Murine leukemia virus
 : US-09-554-572-26

[illegible]

? FILING DATE: 02-AUG-1996
? ATTORNEY/AGENT INFORMATION:
? NAME: Berridge, William P.
? REGISTRATION NUMBER: 30,024
? REFERENCE/DOCKET NUMBER: WPB 38588
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 703-836-6400
? TELEFAX: 703-836-2787
? INFORMATION FOR SEQ ID NO: 87:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 433 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: peptide
US-08-691-563C-87

Query Match	484	DB 3:	Length	433;			
Best Local Similarity	53.08	Pred. No.	9.5e-41;				
Matches	97;	Conservative	41;	Indels	18;	Gaps	4

[illegible]

RESULT 6
 US-09-374-766-87
 : Sequence 87, Application US/09374766
 : Patent No. 6579526
 :
 : GENERAL INFORMATION:
 : APPLICANT: Heve PERRON
 : APPLICANT: Frederic BESEME
 : APPLICANT: Frederic BEDIN
 : APPLICANT: Glaucia PARANHOS-BACCALA
 : APPLICANT: Florence KOMURIAN-PRADEL
 : APPLICANT: Colette JOLIVET
 : APPLICANT: Bernard MANDRAND
 : TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
 : TITLE OF INVENTION: ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLACTIC
 : TITLE OF INVENTION: THERAPEUTIC PURPOSES
 :
 : NUMBER OF SEQUENCES: 92
 :
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Oliff & Berridge
 : STREET: 700 South Washington Street, Suite 300
 : City: Alexandria
 : STATE: Virginia
 : COUNTRY: U.S.A.
 :
 : ZIP: 22314
 :
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patentin Release #1.0, Version #1.30
 :
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/09/374,766
 : FILING DATE:
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US/08/691,563
 : FILING DATE: 02-AUG-1996
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Berridge, William P.

```

?      REGISTRATION NUMBER: 30.024
?      REFERENCE/DOCKET NUMBER: WB 38588
?      TELECOMMUNICATION INFORMATION:
?      TELEPHONE: 703-836-6400
?      TELEFAX: 703-836-2787
?      INFORMATION FOR SEQ ID NO: 87:
?      SEQUENCE CHARACTERISTICS:
?      LENGTH: 433 amino acids
?      TYPE: amino acid
?      TOPOLOGY: linear
?      MOLECULE TYPE: peptide
?      US-09-374-766-87

```

Query Match	16.9%	Score 484	DB 4	Length 433
Best Local Similarity	53.0%	Pred. No. 9.5e-41		
Matches 97	Conservative 27	Mismatches 41	Indels 18	Gaps 4

[illegible]

RESULT 7
 US-08-979-847B-81
 Sequence 81, Application US/08979847B
 Patent No. 6582703
 GENERAL INFORMATION:
 APPLICANT: PERRON, HERVE
 BESEME, FREDERIC
 BEDIN, FREDERIC
 PARANHOS-BACCALA, GLAUCIA
 KOMURIAN-PRADEL, FLORENCE
 JOLIVET-REYNAUD, COLETTE
 MANDRAND, BERNARD
 GARSON, JEREMY
 TURE, PHILIP
 TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
 ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYLACTIC
 THERAPEUTIC PURPOSES
 NUMBER OF SEQUENCES: 210
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: OLIVIER & BERRIDGE, PLC
 STREET: P.O. BOX 19928
 CITY: ALEXANDRIA
 STATE: VA
 COUNTRY: USA
 ZIP: 22320
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/979,847B
 FILING DATE: 26-No. 6582703-1997
 CLASSIFICATION: <Unknown>
 ATTORNEY/AGENT INFORMATION:
 NAME: BERRIDGE, WILLIAM P.
 REGISTRATION NUMBER: 30,024
 REFERENCE/DOCKET NUMBER: WPB 39046A
 TELECOMMUNICATION INFORMATION:

```

REFERENCE/DOCKET NUMBER: WPB 39046A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-836-6400
TELEFAX: 703-836-2787
INFORMATION FOR SEQ ID NO: 109:
SEQUENCE CHARACTERISTICS:
LENGTH: 162 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 109:
US-08-979-847B-109

Query Match      23.4%: Score 672; DB 4; Length 162;
Best Local Similarity 82.1%: Pred. No. 9.5e-61;
Matches 133; Conservative 6; Mismatches 19; Indels 4; Gaps 1;

QY 381 QNRALDLITAEKRGTCFLGEECCYVNSGIVTEKKEIDRIORAEELNTGPMGL 440
    |||||
DB 1 QNRALDLITAKRGSTCLFLGEECCYVNSGIIIEKKEIDRIKCRADLQNTAPMGL 60
    |||||

QY 441 LQWNPWTLPLGLPLAIIILLFGLPCIFNLLVNFVSSRIEAVK---LQMEPKMQSKTK 496
    |||||
DB 61 LQWNPWTLPLGLPLAIIILLFGLPCIFNLLVNFVSSRIEAVKLDIYLOMEPKMQSKTK 120
    |||||

QY 497 IYRRPLDRPASRSDVNDIKGTPPEIISAQPLLRPNAGSS 538
    |||||
DB 121 IYRGPLDRPARLSDVNDIEVTPPEIISTAQPLLRHNSVSS 162
    |||||

RESULT 4
US-08-007-282B-2
Sequence 2, Application US/08007282B
Patent No. 5403582
GENERAL INFORMATION:
APPLICANT: NAZERIAN, KEYVAN
APPLICANT: CALVERT, JAY G.
APPLICANT: WITTER, RICHARD L.
APPLICANT: VANAGIDA, NOBORU
TITLE OF INVENTION: VACCINE COMPRISING FOXLPOX VIRUS
TITLE OF INVENTION: RECOMBINANTS EXPRESSING THE ENVELOPE GLYCOPROTEIN OF AN
TITLE OF INVENTION: AVIAN RETICULOENDOTHELIOSIS RETROVIRUS
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
STREET: 8110 Gatehouse Road Suite 500 East
CITY: Falls Church
STATE: VA
COUNTRY: USA
ZIP: 22042
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/007,282B
FILING DATE: 19930121
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: MURPHY JR., GERALD M.
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 1644-104P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-205-8000
TELEFAX: 703-205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 567 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single

```

```

TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: NO
ANTI-SENSE: NO
US-08-007-282B-2

Query Match      20.7%: Score 593.5; DB 1; Length 567;
Best Local Similarity 33.5%: Pred. No. 8.7e-52;
Matches 154; Conservative 75; Mismatches 176; Indels 55; Gaps 13;

QY 66 TMPRNCYHATLCHNANTHWY-----GKMINPCPGGLVTCVYFT 110
    |||||
DB 121 TMHSTCYEKTOEFTLTKFTYFALLQTKLGSYEDGPKLQASCTGVGRKPCWDPA 180
    |||||

QY 111 QTMGSDGCVQDAEKHYKEVISQLTRVHNGSSPYKGLDSLKHETLRTRLSLFT 170
    |||||
DB 181 PYVSDGGGPTDMIREESVRELEIRHSYPSVOYHPLALPR-----SRGVLDLQ 232
    |||||

QY 171 T---LTGLHEV-SAONP---TNCWICLPLNFRPVSIPVPEQNNFSTEINTSVLVGPL 223
    |||||
DB 233 TSDILAEATHQVLMATNPKLAENCWLCMTLGTIPALP-----TNGNVTLDGNCGLSLPF 287
    |||||

QY 224 VSN---LEIT---HTSNLCVKSNTTYYTNSOCIRWVPTQ-----IVCLPSGIF 269
    |||||
DB 288 GCNPGSIDVSCYAGEADNRGTGPVGYVHF-TNCTSIDEVNETSQMGMLTRLCPPGHV 346
    |||||

QY 270 FVCGTS-AVRCNGSSSESCFLSPVPRMTITYEDD---LYSVVISKPNKR-VPIIPV 324
    |||||
DB 347 FVCGNMMAATYALPNRWICGLCLASIVPDLSITSGEPIPLPIETARHKAQVPIPL 406
    |||||

QY 325 IGAGVIALGTGIGITSTOFPYKLSOELNGDMERVADSLVTLDOULSLAAYVONR 384
    |||||
DB 407 VGLGISGATLAGGTGLGYSVHTYHKLNSQLIDVQALSTINDLQDIDSLAEVLONR 466
    |||||

QY 385 ALDILTAERGTCFLGEECCYVNSGIVTEKKEIDRIORAEELNTGPMGLSQM 444
    |||||
DB 467 GIDLTLAEGGICLALQEKCFYANKSGIVRKIRLDLERRALVDNPLWGLNGF 526
    |||||

QY 445 MPWILPLGLPLAIIILLFGLPCIFNLLVNFVSSRIEAVK 484
    |||||
DB 527 LPYLLPLGLPLGLIILLFGLPCIMKTLRIIHKIQAVK 566
    |||||

RESULT 5
US-08-691-563C-87
Sequence 87, Application US/08691563C
Patent No. 6001987
GENERAL INFORMATION:
APPLICANT: HERVE PERRON
APPLICANT: FREDERIC BESEME
APPLICANT: FLORENCE BESEME
APPLICANT: GLAUCIA PARANHOS-BACCALA
APPLICANT: FLORENCE KOMURIAN-FRABEL
APPLICANT: COLETTE JOLIVET
APPLICANT: BERNARD MANDRAND
TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
TITLE OF INVENTION: ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYL
NUMBER OF SEQUENCES: 92
CORRESPONDENCE ADDRESS:
ADDRESSEE: OLIF & BERRIDGE
STREET: 700 South Washington Street, Suite 300
CITY: Alexandria
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22314
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/691,563C

```

QY 301 TEODLSYVSKPRKRPVILPFVIGAGVLGALGTGIGITTSOIFYKKLSOELNGDME 360
Db 301 TEODLYNVSKPRKRPVILPFVIGAGVLGALGTGIGITTSOIFYKKLSOELNGDME 360
QY 361 VADSLVTLQDOLNSLAAYVQNRALDLTAERGTCFLGEECCYYVNOGIVTEKYKE 420
Db 361 VADSLVTLQDOLNSLAAYVQNRALDLTAERGTCFLGEECCYYVNOGIVTEKYKE 420
QY 421 IRDRIORAEELRNTGPMGLLSOMMPWILPFLGFLAIIILLGPGCIFNLLVNFVSSRI 480
Db 421 IRDRIORAEELRNTGPMGLLSOMMPWILPFLGFLAIIILLGPGCIFNLLVNFVSSRI 480
QY 481 EAVKLOMEPKMQSKTKITRRPLDRPASRSDVNDIKGTPPEISAQPLLRNSAGSS 538
Db 481 EAVKLOMEPKMQSKTKITRRPLDRPASRSDVNDIKGTPPEISAQPLLRNSAGSS 538

RESULT 2
US-08-979-847B-106
; Sequence 106, Application US/08979847B
; Patent No. 6582703
; GENERAL INFORMATION:
; APPLICANT: PERRON, HERVE
; BESEME, FREDERIC
; BEDIN, FREDERIC
; PARANHOS-BACCALA, GLAUCIA
; KOMURIAN-PRADEL, FLORENCE
; JOLIVET-REYNAUD, COLETTE
; MANDRAND, BERNARD
; GARSON, JEREMY
; TUKE, PHILIP
; TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
; ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYL
; THERAPEUTIC PURPOSES
; NUMBER OF SEQUENCES: 210
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OLIEFF & BERRIDGE, PLC
; STREET: P. O. BOX 19928
; CITY: ALEXANDRIA
; STATE: VA
; COUNTRY: USA
; ZIP: 22320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/979, 847B
; FILING DATE: 26-No. 6582703-1997
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: BERRIDGE, WILLIAM P.
; REGISTRATION NUMBER: 30,024
; REFERENCE/DOCKET NUMBER: WPB 39046A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-836-6400
; TELEFAX: 703-836-2787
; INFORMATION FOR SEQ ID NO: 106:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 493 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 106:
US-08-979-847B-106

Query Match 80.6%; Score 2314; DB 4; Length 493;
Best Local Similarity 88.4%; Prcd. No. 2,1e-228;
Matches 433; Conservative 19; Mismatches 38; Indels 0; Gaps 0;

QY 1 MALPHYHFLFTVLLPSTFLTAPPCCRMSSSPYOEFIRMQRPGNIDAPYSLSKGP 60

Db 1 MALPHYHFLFTVLLPSTFLTAPPCCRMSSSPYOEFIRMQRPGNIDAPYSLSKGS 60
QY 61 TFTAHTMPRNCYSATLCMHANTHYWTKMINSFPGGLGVTCWTFYOTGNSDGGV 120
Db 61 TFTAHTMPRNCYSATLCMHANTHYWTKMINSFPGGLGVTCWTFYOTGNSDGGI 120
QY 121 ODAREKIVKEYISQTLVHGTSSPYKGLDLSKIHETLRHTRLVSLFNTLTGLHEVSA 180
Db 121 ODAREKIVKEYISQTLVHGTSSPYKGLDLSKIHETLRHTRLVSLFNTLTGLHEVSA 180
QY 181 QGAREKIVKEYISQTLVHGTSSPYKGLDLSKIHETLRHTRLVSLFNTLTGLHEVSA 180
Db 181 QGAREKIVKEYISQTLVHGTSSPYKGLDLSKIHETLRHTRLVSLFNTLTGLHEVSA 180
QY 241 SNTITTTNSOCIRVYTPPTQIVCLSGIFEVCGTSAYRCLNGSSSMCFSLFVPMY 300
Db 241 SNTITTTNSOCIRVYTPPTQIVCLSGIFEVCGTSAYRCLNGSSSMCFSLFVPMY 300
QY 301 TEODLSYVSKPRKRPVILPFVIGAGVLGALGTGIGITTSOIFYKKLSOELNGDME 360
Db 301 TEODLYNVSKPRKRPVILPFVIGAGVLGALGTGIGITTSOIFYKKLSOELNGDEQ 360
QY 361 VADSLVTLQDOLNSLAAYVQNRALDLTAERGTCFLGEECCYYVNOGIVTEKYKE 420
Db 361 VADSLVTLQDOLNSLAAYVQNRALDLTAERGTCFLGEECCYYVNOGIVTEKYKE 420
QY 421 IRDRIORAEELRNTGPMGLLSOMMPWILPFLGFLAIIILLGPGCIFNLLVNFVSSRI 480
Db 421 IRDRIORAEELRNTGPMGLLSOMMPWILPFLGFLAIIILLGPGCIFNLLVNFVSSRI 480
QY 481 EAVKLOMEPKMQ 492
Db 481 EAVKLOMEPKMQ 492

RESULT 3
US-08-979-847B-109
; Sequence 109, Application US/08979847B
; Patent No. 6582703
; GENERAL INFORMATION:
; APPLICANT: PERRON, HERVE
; BESEME, FREDERIC
; BEDIN, FREDERIC
; PARANHOS-BACCALA, GLAUCIA
; KOMURIAN-PRADEL, FLORENCE
; JOLIVET-REYNAUD, COLETTE
; MANDRAND, BERNARD
; GARSON, JEREMY
; TUKE, PHILIP
; TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
; ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROP
; THERAPEUTIC PURPOSES
; NUMBER OF SEQUENCES: 210
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OLIEFF & BERRIDGE, PLC
; STREET: P. O. BOX 19928
; CITY: ALEXANDRIA
; STATE: VA
; COUNTRY: USA
; ZIP: 22320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/979, 847B
; FILING DATE: 26-No. 6582703-1997
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: BERRIDGE, WILLIAM P.
; REGISTRATION NUMBER: 30,024

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 23, 2003, 14:53:17; Search time 31 Seconds
(without alignments)
734.298 Million cell updates/sec

Title: US-10-069-883-1

Perfect score: 2871

Sequence: 1 MALPFIHFLFTVLLPSFTLT.....PPEISAAQPLRLRNSACSS 538

Scoring table:

BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents-AA:*
- 1: /cgn2_6/ptodata/1/laa/5A.COMB.pep:*
 - 2: /cgn2_6/ptodata/1/laa/5B.COMB.pep:*
 - 3: /cgn2_6/ptodata/1/laa/6A.COMB.pep:*
 - 4: /cgn2_6/ptodata/1/laa/6B.COMB.pep:*
 - 5: /cgn2_6/ptodata/1/laa/PCRTUS.COMB.pep:*
 - 6: /cgn2_6/ptodata/1/laa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2868	99.9	538	4	US-09-175-928-4
2	2314	80.6	493	4	US-08-979-847B-106
3	672	23.4	162	4	US-08-979-847B-109
4	593.5	20.7	567	1	US-08-007-282B-2
5	484	16.9	433	3	US-08-691-563C-87
6	484	16.9	433	4	US-09-374-766-87
7	484	16.9	433	4	US-08-979-847B-81
8	420	14.6	673	3	US-09-075-272-5
9	418.5	14.6	632	4	US-09-315-127-2
10	418.5	14.6	632	4	US-09-315-127-3
11	418.5	14.6	665	4	US-09-309-572-14
12	418.5	14.6	1312	4	US-09-554-572-26
13	412	14.4	192	3	US-08-486-099-107
14	412	14.4	192	3	US-08-360-107A-117
15	412	14.4	192	3	US-08-484-223B-107
16	412	14.4	192	3	US-08-919-597-107
17	412	14.4	192	3	US-08-475-668A-107
18	412	14.4	192	3	US-08-485-551A-107
19	412	14.4	192	3	US-08-471-913A-107
20	412	14.4	192	3	US-08-485-264A-107
21	412	14.4	192	4	US-08-474-349A-107
22	412	14.4	192	4	US-08-470-896-107
23	412	14.4	192	4	US-08-485-546A-107
24	411.5	14.3	638	3	US-09-376-781-6
25	405.5	14.1	654	4	US-09-315-127-11
26	405.5	14.1	654	4	US-09-315-127-12
27	403.5	14.1	660	3	US-09-111-085-2

28	403.5	14.1	660	3	US-09-376-781-5	Sequence 5, Appl1
29	400.5	13.9	645	4	US-09-315-127-8	Sequence 8, Appl1
30	400.5	13.9	645	4	US-09-315-127-9	Sequence 9, Appl1
31	390.5	13.6	657	3	US-09-111-085-4	Sequence 4, Appl1
32	390.5	13.6	657	3	US-09-376-781-4	Sequence 4, Appl1
33	389.5	13.6	667	4	US-09-315-127-5	Sequence 5, Appl1
34	389.5	13.6	667	4	US-09-315-127-6	Sequence 6, Appl1
35	389	13.5	232	4	US-09-555-352-10	Sequence 10, Appl1
36	386.5	13.5	656	3	US-09-376-781-3	Sequence 3, Appl1
37	370	12.9	67	1	US-08-686-878A-49	Sequence 49, Appl1
38	370	12.9	67	1	US-08-721-489-3	Sequence 3, Appl1
39	323	11.3	196	2	US-08-484-126-7	Sequence 7, Appl1
40	323	11.3	196	4	US-09-374-909-7	Sequence 7, Appl1
41	291	10.1	152	4	US-09-120-653D-7	Sequence 7, Appl1
42	287	10.0	77	4	US-08-979-847B-103	Sequence 103, Appl1
43	285.5	9.9	227	4	US-09-904-615-126	Sequence 126, App
44	213.5	7.4	486	1	US-07-672-483-2	Sequence 2, Appl1
45	210	7.3	486	4	US-08-259-451-13	Sequence 13, Appl1

ALIGNMENTS

RESULT 1
US-09-175-928-4
Sequence 4, Application US/09175928A

Patent No. 6312921

GENERAL INFORMATION:

APPLICANT: Jacobs, Kenneth

APPLICANT: McCoy, John M.

APPLICANT: Lavallie, Edward R.

APPLICANT: Collins-Racle, Lisa A.

APPLICANT: Evans, Cheryl

APPLICANT: Merberg, David

APPLICANT: Treacy, Maurice

APPLICANT: MI, Sha

APPLICANT: Genetics Institute, Inc.

TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM

FILE REFERENCE: 6006B A1172A

CURRENT APPLICATION NUMBER: US/09/175,928A

CURRENT FILING DATE: 1998-10-20

NUMBER OF SEQ ID NOS: 62

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 4

LENGTH: 538

TYPE: PRT

ORGANISM: Homo sapiens

US-09-175-928-4

Query Match 99.9%; Score 2868; DB 4; Length 538;

Best local similarity 99.8%; Pred. No. 3.5e-285;

Matches 537; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY	1	MALPFIHFLFTVLLPSFTLTAPPCRCMTSSPYQEFRLMKORPGNIDAPSTRSLSKTP	60
DB	1	MALPFIHFLFTVLLPSFTLTAPPCRCMTSSPYQEFRLMKORPGNIDAPSTRSLSKTP	60
QY	61	TFTAHTHMPRNCYHSATLCMHANTHTYMTGKMINPSCPGGLGTVTCVTFQTQMSDGGGV	120
DB	61	TFTAHTHMPRNCYHSATLCMHANTHTYMTGKMINPSCPGGLGTVTCVTFQTQMSDGGGV	120
QY	121	ODQAREKHKEYISQLTRVHGTSPPYKGLSLKHLHTLTHRLVSLFTTGLHEVSA	180
DB	121	ODQAREKHKEYISQLTRVHGTSPPYKGLSLKHLHTLTHRLVSLFTTGLHEVSA	180
QY	181	ONPTNCWICLPLNFRPYVSIYPVEOMNNTSTELNTTSVYGLVSLULETHHSNLTGVCF	240
DB	181	ONPTNCWICLPLNFRPYVSIYPVEOMNNTSTELNTTSVYGLVSLULETHHSNLTGVCF	240
QY	241	SNNTYTTNSOCLRWVPPPOIYCLPSGIFVCGTSAYRCLNNGSSBESMCFSLFVPPMTTY	300
DB	241	SNNTYTTNSOCLRWVPPPOIYCLPSGIFVCGTSAYRCLNNGSSBESMCFSLFVPPMTTY	300

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 23, 2003, 14:52:53 ; Search time 43 Seconds
(without alignments)
1203.227 Million cell updates/sec

Title: US-10-069-883-1

Perfect score: 2871

Sequence: 1 MALPYHIFLEFVLLPSFTLT.....PPEEISAQPLLRNSAGSS 538

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	627	21.8	563	1 VCMVM7	env polypeptin -
2	599	20.9	587	1 VCLJSA	env polypeptin -
3	593.5	20.7	567	1 VCFVAS	env polypeptin -
4	593.5	20.7	586	1 VCLJMP	env polypeptin -
5	549	19.1	575	1 VCLJHD	env polypeptin -
6	508.5	17.7	570	4 B44282	retrovirus-related
7	502	17.5	582	1 VCVDA8	env polypeptin -
8	465.5	16.2	321	2 S12812	env polypeptin -
9	424	14.8	640	1 VCMVRV	env polypeptin -
10	419	14.6	676	2 T01381	env polypeptin -
11	418.5	14.4	665	1 VCMVEM	env polypeptin -
12	414	14.4	676	2 S70395	env polypeptin -
13	413.5	14.4	688	2 A43491	env polypeptin -
14	413.5	14.4	688	2 B43491	env polypeptin -
15	412.5	14.4	662	1 VCMVGF	env polypeptin -
16	412.5	14.4	662	1 VCMVGL	env polypeptin -
17	411.5	14.3	662	1 VCMVLA	env polypeptin -
18	410	14.3	676	1 VCMVPA	env polypeptin -
19	409	14.2	645	1 VCMVSS	env polypeptin -
20	408	14.2	668	1 VCMVFP	env polypeptin -
21	407.5	14.2	636	1 VCMVFS	env polypeptin -
22	406.5	14.2	636	1 VCMVFS	env polypeptin -
23	406.5	14.2	669	2 T10533	env polypeptin -
24	406.5	14.2	669	2 VCMVEM	env polypeptin -
25	406	14.1	627	1 VCMVM2	env polypeptin -
26	403.5	14.1	642	1 VCMVFG	env polypeptin -
27	402	14.0	640	1 VCMVM1	env polypeptin -
28	401.5	14.0	669	2 A46511	envelope protein -
29	401	14.0	665	1 VCMVVR	env polypeptin -

30	399.5	13.9	639	1 VCMVSA	env polypeptin pr
31	399	13.9	665	1 VCMVKA	env polypeptin pr
32	389.5	13.6	644	2 S15464	env polypeptin - mtr
33	389.5	13.6	667	1 VCLJGL	env polypeptin pr
34	387.5	13.5	661	1 VCMVCB	env polypeptin -
35	386.5	13.5	353	2 S22805	env polypeptin -
36	384.5	13.4	666	1 VCMVHL	env polypeptin -
37	373	13.0	671	1 VCMVCE	env polypeptin -
38	330.5	11.5	211	2 A03986	env polypeptin -
39	327.5	11.4	201	2 A03985	env polypeptin -
40	213.5	7.4	486	1 VCLJH2	env polypeptin -
41	213.5	7.4	486	1 VCLJH2	env polypeptin -
42	209.5	7.3	247	2 A53482	transmembrane enve
43	209	7.3	247	2 A53482	transmembrane enve
44	206	7.2	488	1 VCLJCN	env polypeptin -
45	203	7.1	488	1 VCLJMT	env polypeptin -

ALIGNMENTS

RESULT 1

VCMVM7 env polypeptin - baboon endogenous virus (strain M7)

N:Alternate names: coat polypeptin

N:Contains: coat protein gp70; coat protein p20E

C:Species: baboon endogenous virus

A:Note: host Papio sp. (baboon)

C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 16-Jul-1999

C:Accession: J02622

R:Kato, S.; Matsuo, K.; Nishimura, N.; Takahashi, N.; Takano, T.

A:Title: The entire nucleotide sequence of baboon endogenous virus DNA: a chimeric ge

A:Reference number: J02622

A:Accession: J02622

A:Molecule type: DNA

A:Residues: 1-563 <KAT>

A:Cross-references: GB:ML6550; NID:9509586; PIDN:AA07333.1; PID:9332599

C:Genetics:

A:Gene: env

C:Superfamily: type C retrovirus env polypeptin

C:Keywords: coat protein; glycoprotein; polypeptin; transmembrane protein

F:1-376/Product: coat protein gp70 #status predicted <CGP>

F:377-563/Product: coat protein p20E #status predicted <CGP>

F:437-449/Region: immunosuppressive peptide #status predicted

F:113,219,228,264,282,292,306,312,321,339,469/Binding site: carbohydrate (Asn) (covar

Query Match 21.8%; Score 627; DB 1; Length 563;

Best local similarity 31.6%; Pred. No. 5.8e-39;

Matches 158; Conservative 87; Mismatches 169; Indels 86; Gaps 16;

QY	31	SSPYQEFIMRMORPCGNIDAPSTRSLSKGTPPTTAHTMRNCYHSATLCHMANTHYW---	87
DB	83	SSPQDEC-----PCN-----SVHS-----SCYTSYQCRSGNKRYTAT	120
QY	88	-----TGKMINPSCPGIGLVTCVCTYTGMSGSGGVDOARCKHYK	130
DB	121	LKQTGTGSDVQVLASTKLIQSPCNGIKGOSICSTAPRHVSOGGGLDITR---IK	177
QY	131	EVISQITRVHGTSSP---YKGLDSKLAEHTLRTLRVLSVFTTTLGLHEVSQNPNTCW	187
DB	178	SVQRKLEIEHKALYELQYHPLAIPKVRDNLMVDAQTLINATYVLLMLMSNTSLVDDCW	237
QY	188	ICLPLNFRPVVISIPPEQNNNSTEL--NTTSVYQPLVLSNLEITTSULTC-----	237
DB	238	ICLPLNFRPVVISIPPEQNNNSTEL--NTTSVYQPLVLSNLEITTSULTC-----	237
QY	238	-----VKFSNTYTTNSQCIRWVPTPTQVLCPSGIFVCGTS--AVRCLNGSSSNC	288
DB	294	TEIDIGHAVAFSNCSTITN-----VTGP---ICAVNGSVFLCGNMATYTLPTNNTGIC	344
QY	289	FLSFLVPPMTIYTEDQ-----LTSYVYSRPRNKRVPILPEVIGAGVLGALGTCIGGITTSS	343

Db 527 LPYLLPLGLFLGLILFLTLGPGIMKTLRIIHDRIOAVK 566

RESULT 4

VCLJMP
env polypolypeptide - Mason-Pfizer monkey virus
N:Alternate names: coat polypolypeptide
M:Contains: coat protein gp20; coat protein gp70
C:Species: Mason-Pfizer monkey virus
C:Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 16-Jul-1999
C:Accession: D25839
R:Sonigo, P.; Barker, C.; Hunter, E.; Wain-Hobson, S.
Cell 45, 375-385, 1986
A:Title: Nucleotide sequence of Mason-Pfizer monkey virus: an immunosuppressive D-type
A:Reference number: A90878; MUID:86189951; PMID:2421920
A:Accession: D25839
A:Molecule type: DNA
A:Residues: 1-586 <SON>
A:Cross-references: GB:M2349; NID:9334702; PIDN:AAA7712.1; PID:9334705
A:Experimental source: clone 6A
C:Genetics:
A:Gene: env
C:Superfamily: type C retrovirus env polypolypeptide
C:Keywords: capsid protein; coat protein; glycoprotein; polypolypeptide
F:1-20/Domain: signal sequence #status predicted <Sig>
F:21-394/Product: coat protein gp70 #status predicted <GP7>
F:395-586/Product: coat protein gp20 #status predicted <GP2>
F:120,237,264,276,291,304,318,324,339,357,487/Binding site: carbohydrate (Asn) (covalent

Query Match 20.7%; Score 593.5; DB 1; Length 586;
Best Local Similarity 28.6%; Pred. No. 1.9e-36;
Matches 156; Conservative 89; Mismatches 192; Indels 109; Gaps 16;

24 PCRCTSSPYGEFLMRQRPENIDAPYSRLSK-----GTPPT 63
41 PCDC-----AGSYSSPPIINSLLTSCSTHAYSTNSLMKOCVSTPTTP 85
64 ANTH-----MPRNCYHATLCMHANTHYWG----- 89
86 SMTHGSCPGECNITSYDVHASCYNHQCGNIGKTYLATITGDRPAIDGCVPTVL 145
90 ----KMINSPPG-LGYTCWTFYTGMSDGGVQDQAREKHKEVISQLTRHGTSS 144
146 GTHNLITAGCPNGKKGVCWNSRPSVHISDGGPQDKARDIYKKFEELHRSLEPFL 205
145 PKKGLDLSKLH-ELTRHT-----RLVSLFNTITLGLHEVSAQNPTNCWICLPNFRY 197
206 SHPLALPARGKEKIDAHITDLDLATVHSLNASSQPSLAE-----DCMLCIGSDPVP 258
198 VSIPEPEO-WNNFSTEINTTAVLYGP-LVSNLEITHTSNLCVCFESNTTYTN-----S 249
259 LALPNTDLCNFAALSHNSCLTPPLVQPNFT-DSCLTAHYNNSFDIDVGLAST 317
250 QC-----IRWTPPTQIVCLPSGIFVCGTS-AYRCLNGSSSESMCFSLVPPMTITYEO 303
318 NCSSYYNASTASKPNSICAPNSVFCGNNKATYLLPMTGSCVIALLLDIDIPES 377
304 D-----LVSYSKRNRKRVPLIPVIGAGVIGALGTGTTSTQFYKYSOELNGCM 358
378 EEPVPIALDHLGKAK-RAIDQILPFGITTAAGTGAAGVISTQYTKLSHOLISRY 436
359 ERVADSLVTLQDOLSLAAVYVONRRALDLLTAEGGCTCLFGECCYVYVNSGIVTEKV 418
437 QAIISTIDLDQVSLAEVYVQNRGDLTLAEGGCTCLAOECCCFANSGIVRDT 496
419 KEIRDRIQRAEELNNTGPMGLISQMPWILPPLGLAIIILLFGPCIFNLVNFVS 478
497 KMLODLERRRQLDNPFTSFGFLPVPMLDPLCLLLVLSFGPIIFKMLTFIKH 556
479 RTEAVK 484
557 QIESIO 562

RESULT 5

VCLJHD
env polypolypeptide precursor - squirrel monkey retrovirus SMRV-H
N:Alternate names: coat polypolypeptide
M:Contains: coat protein gp20; outer membrane protein
C:Species: squirrel monkey retrovirus SMRV-H
C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 16-Jul-1999
C:Accession: D31827
R:Oda, T.; Ikeda, S.; Watanabe, S.; Hattushike, M.; Akiyama, K.; Mitsunobu, F.
Virology 167, 468-476, 1988
A:Title: Molecular cloning, complete nucleotide sequence, and gene structure of the p
A:Reference number: A31827; MUID:89073750; PMID:3201749
A:Accession: D31827
A:Molecule type: DNA
A:Residues: 1-575 <ODA>
A:Cross-references: GB:M23385; NID:9332626; PIDN:AAA66455.1; PID:9332627
C:Genetics:
A:Gene: env
C:Superfamily: type C retrovirus env polypolypeptide
C:Keywords: capsid protein; coat protein; glycoprotein; polypolypeptide; transmembrane pr
F:1-17/Domain: signal sequence #status predicted <Sig>
F:18-386/Product: outer membrane protein #status predicted <OMP>
F:387-575/Product: coat protein gp20 #status predicted <TMP>
F:387-403/Domain: transmembrane #status predicted <TM1>
F:447-481/Region: immunosuppressive peptide
F:518-534/Domain: transmembrane #status predicted <TM2>
F:126,239,266,271,302,316,322,349,479/Binding site: carbohydrate (Asn) (covalent) #st

Query Match 19.1%; Score 549; DB 1; Length 575;
Best Local Similarity 30.6%; Pred. No. 4e-33;
Matches 148; Conservative 83; Mismatches 148; Indels 105; Gaps 18;

66 THMPRNCYHATLCMHANTHYWGK-----INP-----SCPGLGATV 104
109 SOMHSSCYSSFSQCCQGNNTYFTALIQRTKSTETNPSGLQPHGLQAGCDGVGKSV 168
105 CWTYFQTCMSDGGVQDQAREKHKEV-----SOLTRVH-----GTSSPYKGLDLSKLH 155
169 CMMQQAPTIVSDGGPQDQVRELIVQKQIELVYIOSQFRLSHPLARSKP-RGPDID 224
156 ETLRHTRLVSLFNTITLGLHEVSAQNPTNCWICL-----PLNF----- 194
225 -----AQMLDILSATHQALNINSPSLAQCWMLCINQGSMLAPPVNISFNAGQNNCT 278
195 --RPVYSLPPEO-WNNF--STEINTSVLQPLVSNLEITHTSNLTCKESNTTYTNS 249
279 PSLPFRVQPMPSQVPCFFKGAQNNSEFDIPG--VAN-----FVNCSSSNH 323
250 QCIIRWTPPTQIVCLPSGIFVCGTS-AYRCLNGSSSESMCFSLVPPMTITYEOD 304
324 -----SEALCPEPGCAFVCGNNLAETJALPAMNTGSCVIALLLDIDISGDDPVI 374
305 -LVSYSKRNRKRVPLIPVIGAGVIGALGTGTTSTQFYKYSOELNGCMERVD 363
375 PTFEDYIAGR-QKRAVTLITLVLGLGVSTVAATGACGVAVSYTKLSQLLINDVALSS 433
364 SLVTLQDQINSLAAVYVONRRALDLLTAEGGCTCLFGECCYVYVNSGIVTEKV 423
434 TINDLDQDLSLAEVYVQNRGDLTLAEGGCTCLAOECCCFANSGIVRDT 493
424 RIQR-----AEELNNTGPMGLISQMPWILPPLGLAIIILLFGPCIFNLVNFVS 479
494 DLEKRRKALADNMLFTGLNGL----PYLLPGLPGLAIIILFFSFAPWILRRVATLIRQ 549
480 TEAV 483
550 LMSL 553

RESULT 6

B44282


```

Db      548  LQNRGDLILFLKEGGLCAALKECCFYADHTGLVRDSMAKLREHLSQKQLFESSQW- 606
Qy      440  LLSQMW---PW-----LLPELPLAAILLLFGPCTILYNFYSRIEANK----- 484
Db      607  -PEGMFNRSWFTLLISTINGPLILILLFLGPGCLNRLVQFYADRSIQVALVLTQOY 665
Qy      485  -----LMEPK 490
Db      666  HQKPLEVEPO 676

```

RESULT 11
VCVWEM

env polypeptide - Moloney murine leukemia virus
N:Contents: knob protein gp70; R protein; spike protein p15E
C:Species: Moloney murine leukemia virus
C:Date: 01-Sep-1981 #sequence_revision 24-Sep-1981 #text_change 16-Jul-1999
C:Accession: A93265; A93235; A93848; A03983
R:Shinnick, T.M.; Lerner, R.A.; Sutcliffe, J.G.
N:ature 293, 543-548, 1981
A:Title: Nucleotide sequence of Moloney murine leukemia virus.
A:Reference number: A93265; MUID:82035843; PMID:6169994
A:Accession: A93265
A:Molecule type: genomic RNA
A:Residues: 1-665 <SH1>
A:Cross-references: GB:J02255; GB:J02256; GB:J02257; GB:M76668; NID:g331934; PIDN:AAB5995
A:Experimental source: clone pMLV-1
R:Sutcliffe, J.G.; Shinnick, T.M.; Green, N.; Liu, F.T.; Niman, H.L.; Lerner, R.A.
N:ature 287, 801-805, 1980
A:Title: Chemical synthesis of a polypeptide predicted from nucleotide sequence allows d
A:Reference number: A93235; MUID:81052384; PMID:6159543
A:Accession: A93235
A:Molecule type: genomic RNA
A:Residues: 496-665 <SUT1>
A:Experimental source: provirus
R:Sutcliffe, J.G.; Shinnick, T.M.; Verma, I.M.; Lerner, R.A.
Proc. Natl. Acad. Sci. U.S.A. 77, 3302-3306, 1980
A:Title: Nucleotide sequence of Moloney leukemia virus: 3' end reveals details of replid
A:Reference number: A93848; MUID:81013872; PMID:6251454
A:Accession: A93848
A:Molecule type: DNA
A:Residues: 484-662; 'CEP' <SUT2>
A:Experimental source: provirus, clone pMLV-201
C:Genetics:
A:Gene: env
C:Superfamily: type C retrovirus env polypeptide
C:Keywords: coat protein; glycoprotein; polypeptide; spike protein; transmembrane protei
F:34-469/Product: knob protein gp70 #status predicted <KN>
F:470-649/Product: spike protein p15E #status predicted <SPK>
F:650-665/Product: R protein #status predicted <RP>
F:45-199,326,358,398,434/Binding site: carbohydrate (Asn) (covalent) #status predicted

```

Db      ||||| ||| : ||| : : : : : : : : : : : : : : ||
538  LQNRGIDLLFLKEGGJCAALKEECFCYADHGLVRDSMAKRETLNQRKLFSTGCMF 597

Oy      439  -GLTSQMPW----ILPGLCAPIAIIILFGPCIFNLVNVSSRIABVK 484
      || : || : : || : : : : : : : : : : || : || :
Db      598  ECLFNR-SPEFTLLISTIMGPLVILMIILFGPCILNKLRYQVFKRISVQ 647

```

RESULT 12

env polypeptide - Friend murine leukemia virus (strain FB29)
N:Contents: knob protein gp75; R protein; spike protein p15E
C:Species: Friend murine leukemia virus
A:Variety: strain FB29
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 20-Sep-1999
C:Accession: S70395
R:Perryman, S.; Mishlo, J.; Chesebro, B.
Nucleic Acids Res. 19, 6950, 1991
A:Title: Complete nucleotide sequence of Friend murine leukemia virus, strain FB29.
A:Reference number: S70393; MUID:92107687; PMID:1762923
A:Accession: S70395
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: genomic RNA
A:Residues: 1-676 <P>E</P>
A:Cross-references: EMBL:211128; NID:961547; PIDN:CAA77479.1; PID:961549
A:Experimental source: strain FB29
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1991
C:Genetics:
A:Gene: env
C:Superfamily: type C retrovirus env polypeptide
C:Keywords: coat protein; polypeptide; spike protein; transmembrane protein

RESULT 13
A43491

env polypeptide-1 (B-7851/B murine leukemia virus (strain T1223/B)
C:Species: B-7851/B murine leukemia virus
C:Date: 12-Jan-1993 #sequence_revision 12-Jan-1993 #text_change 20-Apr-2001
C:Accession: A43491
R:Baylac-Kalabokias, H.; Astler-Gin, T.; Moynet, D.; Hermould, M.; Mamoun, R.; Legran
Virus Res. 18, 117-134, 1991

```

Db      382 -PANCSTVASQHKITL-----SEVTGCGICIGAVPKTIOWINNTQKSTGCSYIARDGCM 436
          :| : || : || || :|| :|

```

[illegible]

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 23, 2003, 14:45:22; Search time 26 Seconds

(without alignments)
973.091 Million cell updates/sec

Title: US-10-069-883-1

Perfect score: 2871

Sequence: 1 MALPYHIFLVLPSTLT.....PPETISAAQPLLRPSAGSS 538

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	627	21.8	563	1 ENV_BAEVM	P10269 baboon endo
2	620.5	21.6	574	1 ENV_SNV2	P51515 simian retr
3	601.5	21.0	574	1 ENV_SNV2R	P51520 simian retr
4	599	20.9	587	1 ENV_SNV1	P04027 simian retr
5	593.5	20.7	567	1 ENV_AVISN	P31796 avian splice
6	593.5	20.7	586	1 ENV_MPMV	P07575 simian maso
7	549	19.1	575	1 ENV_SMRVH	P21412 squirrel mo
8	502	17.5	582	1 ENV_AVIRE	P03399 avian relic
9	424	14.8	640	1 ENV_RMCVF	P06445 rauscher m
10	418.5	14.6	665	1 ENV_MLVMO	P26804 moloney mur
11	414	14.4	676	1 ENV_MLVVF	P26804 friend mur
12	412.5	14.4	662	1 ENV_PSVGA	P03390 feline sarc
13	412	14.4	675	1 ENV_MLVFS	P03390 feline sarc
14	411.5	14.3	662	1 ENV_PVLVB	P11261 feline leuk
15	410	14.3	676	1 ENV_MLVFP	P26803 feline leuk
16	409.5	14.3	642	1 ENV_PVLGL	P08359 feline leuk
17	409	14.2	665	1 ENV_FSVSM	P21445 feline sarc
18	408	14.2	668	1 ENV_FLVCS	P21443 feline leuk
19	406.5	14.2	636	1 ENV_MCFP	P15073 mink cell f
20	406.5	14.2	669	1 ENV_MLVAV	P03386 akv murine
21	402	14.0	640	1 ENV_MCFV3	P03386 akv murine
22	401	14.0	665	1 ENV_MLVSD	P11268 radiation m
23	399.5	13.9	639	1 ENV_FLYSA	P06752 feline leuk
24	399.5	13.9	679	1 ENV2_MOUSE	P11370 mus musculu
25	399	13.9	665	1 ENV_MLVYK	P31794 radiation m
26	389.5	13.6	667	1 ENV_GALV	P21415 gibdon ape
27	387.5	13.5	661	1 ENV_MLVCB	P08360 cas-br-e mu
28	384.5	13.4	666	1 ENV_MLVHO	P21436 hominyl mur
29	373	13.0	671	1 ENV_FENV1	P31791 feline endo
30	330.5	11.5	211	1 ENV_MSVFV	P03387 fbj murine
31	327.5	11.4	201	1 ENV_MLVKI	P04502 kitsten mur
32	213.5	7.4	486	1 ENV_HTLV2	P03383 human t-cel
33	206	7.2	488	1 ENV_HTLVC	P14075 human t-cel

34	203	7.1	488	1 ENV_HTLIM	P23064 human t-cel
35	201	7.0	488	1 ENV_HTLIF	P03817 human t-cel
36	201	7.0	488	1 ENV_HTLIN	P03816 human t-cel
37	187.5	6.5	488	1 ENV_HTLIA	P03381 human t-cel
38	176	6.1	534	1 ENV_FSVST	P03392 feline sarc
39	167.5	5.8	515	1 ENV_BLVAV	P25057 bovine leuk
40	163.5	5.7	515	1 ENV_BLV	P51519 bovine leuk
41	159.5	5.6	515	1 ENV_BLVBS	P25507 bovine leuk
42	157.5	5.5	515	1 ENV_BLVAF	P25504 bovine leuk
43	157.5	5.5	515	1 ENV_BLVAV	P25505 bovine leuk
44	153.5	5.3	515	1 ENV_BLVJ	P03380 bovine leuk
45	152	5.3	676	1 VCP_EBOIC	P06810 ebola virus

ALIGNMENTS

RESULT 1

ID ENV_BAEVM STANDARD: PRT: 563 AA.

AC P10269; 01-MAR-1989 (Rel. 10, Created)

DT 01-MAR-1989 (Rel. 10, Last sequence update)

DE 16-OCT-2001 (Rel. 40, Last annotation update)

DE ENV polypeptide precursor (Coat polypeptide) [contains: Outer membrane protein GP70; Transmembrane protein P20E].

GN ENV.

OS Baboon endogenous virus (strain M7).

OC Viruses; Retroviral viruses; Retroviridae; Mammalian type C retroviruses.

OX NCBI_TaxID=11764;

RN [1]

RP Kato S., Matsuo K., Nishimura N., Takahashi N., Takano T.;

RA "The entire nucleotide sequence of baboon endogenous virus DNA: a

RT chimeric genome structure of murine type C and simian type D

RT retroviruses.";

RL Jpn. J. Genet. 62:127-137(1987).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC EMBL: D10032; BAA00924.1; -

DR EMBL: X05470; CAA29028.1; -

DR PIR: JTD262; VCMVW7.

DR HSSP: P03385; IMOF.

DR Interpro: IPR002050; Env_polypeptide.

DR Pfam: PF00429; ENV_polypeptide.1.

KW Coat protein; Glycoprotein; Polypeptide; Transmembrane; Signal.

KW SIGNAL

FT CHAIN	1	346	OUTER MEMBRANE PROTEIN GP70 (POTENTIAL).
FT	21	563	TRANS-MEMBRANE PROTEIN P20E (POTENTIAL).
FT CARBOHYD	113	113	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT CARBOHYD	219	219	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT CARBOHYD	229	229	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT CARBOHYD	264	264	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT CARBOHYD	282	282	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT CARBOHYD	292	292	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT CARBOHYD	306	306	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT CARBOHYD	312	312	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT CARBOHYD	321	321	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT CARBOHYD	339	339	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT CARBOHYD	469	469	N-LINKED (GLCNAc. . .) (POTENTIAL).
SO SEQUENCE	563 AA;	61879 MW;	9573137DCA620BB7 CRC64;

Query Match Best Local Similarity 21.8%; Score 627; DB 1; Length 563; Matches 158; Conservative 87; Mismatches 169; Indels 86; Gaps 16;

```

OY 31 SSPVOEFLMRQRPCNIDAPSVRSLSKGPRTTAAHMRNRNCXSLTCLMNAHTYH--- 87
Db 83 SGPIQEC-----PCN---STOS-----SVHS---SCTSTYQOORSGKTYTAT 120
OY 88 -----TGKMINPSCPGGLGYVCVCTYPTOTGMSDGGGVODAREKHVK 130
Db 121 LLKTGTGSDVQVIGSTNKLTIQSPCNCKIGQSIGCWSTAPJHVSOGGGLDTR--IK 177
OY 131 EVISGLTVHGHSSP---YKGLDSKLHETLKHTRVLSFNTTLTGLGHEVSAQNPTNC 187
Db 178 SVQRLEEHIALYELQHPHAIKRVKRNLMVDAQFLMLNATYLLMLMSNTSLVDWC 237
OY 188 ICLPLNFRPYVASIPPEQOMNPNSTET--NTTSVLVQPLSNLEITHTSNLTC----- 237
Db 238 LCLKRG--PPLPLAIKPNFLSVTHSSDNISCLIIIPPL--VQPMFMSNSJCIFRPSYNS 293
OY 238 -----YKRSNTTYTTNSOCIRWTPPTQIVCLPSGFIYVCGIS--AYRCLNGSSESMC 288
Db 294 TEEDIDLGVAHFAESCNCSTITN-----VTGJ--ICANGSVFLCNGMNAATYLETNNWTGIC 344
OY 289 FLSFLVPMPTIYTEOD-----LYSVYISKRPNKRNVILPFVIGAGVGLAGTIGGITTS 343
Db 345 VLATLLPIDIDILGGEPAVPIRAIDHPIYRPK-RALQFIPLLAGLITTAFTTGAITGLGS 403
OY 344 TOFYKKLSQELNGDMERVAOSLVTTIQQDNLNSLAAYVLONRBALDLTAEBGTCFLVGE 403
Db 404 VTQYTKLSNQLISDVOILSTTQIDLODQVDSLAEVVLQNRGIDLLTAEQGGICLALQEK 463
OY 404 CCYYVNOGSIYVEKVEIKEDRIQRAAEYELNTPGMCILSOMPMWILPFLQPLAIIILL 463
Db 464 CCFVYNKSGIYADKIKTLQDELEERRRDKDLASNPMTGLOGCLPYLPLPFLGPIILLTLL 523
OY 464 GPCCLFNLLNVFSSRIEAV 483
Db 524 IGPCIFNRLTAEINDKLNII 543

```

```

RESULT 2
ENV_SRV2
ID ENV_SRV2 STANDARD: PRT: 574 AA.
AC P51515.
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE ENV polyprotein (Coat polyprotein) [Contains: Coat protein GP70; Coat
  protein GP20].
GN ENV.
OS Simian retrovirus SRV-2.
OC Viruses: Retrovirdae; Retroviridae; Betaretrovirus.
OX NCBI_TaxID=39068;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87151131; PubMed=2435057;
RA Thayer R.M., Power M.D., Bryant M.L., Gardner M.B., Barr P.J.,
  Luciw P.A.;
RT "Sequence relationships of type D retroviruses which cause simian
  acquired immunodeficiency syndrome.";
RL Virology 157:317-328(1987).
CC -I- PRT: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: M16605; AAA47563.1; ALT_INT.
DR HSSP: P03385; IMOF.
DR InterPro: IPR002050; Env_polyprotein.
DR Pfam: PF00429; Env_polyprotein; 1.
DR Coet protein; glycoprotein; Polyprotein; Transmembrane.

```

FT	CHAIN	1	382	COAT PROTEIN GP70.
FT	CHAIN	383	574	COAT PROTEIN GP20.
FT	TRANSMEM	386	406	POTENTIAL.
FT	TRANSMEM	515	535	POTENTIAL.
FT	CARBOHYD	117	117	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	233	233	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	260	260	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	267	267	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	288	288	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	298	298	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	312	312	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	318	318	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	327	327	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	345	345	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	475	475	N-LINKED (GLCNAC. . .) (POTENTIAL).
SO	SEQUENCE	574 AA:	62952 MM:	662B1E11A437E055 CRC64.
Query Match				
Best Local Similarity		21.6%: Score 620.5; DB 1; Length 574;		
Matches 153; Conservative		94; Mismatches 195; Indels 65; Gaps 14		

```

0Y 25 GCMCMSSSSYYOEFELRMORPCNIDAPSYLSKSG-----DPTFAAHMHPNCKCHSAT 77
Db 62 CSSHAPOPSDLSKWR-----CVSNPTLANGENIGNCPOCKTFKESVH--SSCYTAYQ 111
0Y 78 LCMHANTHYMG-----KMINSPCGJGVYWCMTYFTOTGH 114
Db 112 ECFEGNKYITYTALILASNRAPITGTSNPTVLNCTNHNLSAGCTGNGVOPICMPKPAVHI 171
0Y 115 SDGGGVODQAREKHVEVYSOLTFRVCHTSSP--YKGLDLSKHETLRTRHLSLENTT 171
Db 172 SDGGGVODQAREKIAVOK--RLBEIHKLSLEPRLRYPLRLPARCKEX IDAQFNNILTAT 228
0Y 172 LTGLHEVASQNTKNCWICLPLNFRPVSLPVEDQMNSTETINTTSLVLP--LYSNLEIT 230
Db 229 YSLNKSXPNLANECWMLCP--SGNP--IPLAIPNSDSFLGSLSCPILPRLVQGLEPMA 284
0Y 231 HTSNLTGY--KFSNTTYTTNSOCIRMWVPPPOI----VCLPSGJFEVCGTS--AYRCLNG 282
Db 285 NLINASCSYSPQNNNSFOVDVGLVERANCSSTLTINSHSLCARSSSFVCGNKNKATYYLPS 344
0Y 283 SSESNCETSLVLPMTIY----TEODLYSYVISKPRNKRVPILPFVIGAVLGALGTGI 337
Db 345 NMTGCVLATLILPDIDIVGDAVPVPAIDHYLHRAR--RAYOFILPVLGIGITTAVSTGT 403
0Y 348 GGITTSQPFYKLSOGLNOMDERVADSLYTLODONSILAUVYONERALDILTAEGGNC 397
Db 404 AGLGYSTOYTKLSLQSLSDVAIISTIDDLODODVSLAEVYONNRGIDLLTAEDGJC 463
0Y 398 FLFEGECYYVWOSGIVTEKVEKEIRDIORAEELNNTGPMGLISQMPMILPFLGLPLAA 457
Db 464 LALQKCCFYNKSGSIVRDKIKRLQDEDEKRRKEIIDNPFWGLHGLLYLLPLGLPPLC 523
0Y 458 ILLLLFGPCIFNLLVNFYSRIEAVK 484
Db 524 LLLLTTFGPLIFPNKIITFEVKGQIDAQ 550

```

```

RESULT 3
ENV_SrvR2
ID ENV_SrvR2 STANDARD: PRT: 574 AA.
AC p51520;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE ENV polypeptide (Coat polypeptide) [Contains: Coat protein GP70; Coat
DE protein GP20].
GN ENV.
OS Simian retrovirus SRV-2 (isolate 2R-18B1).
OC Viruses; Retrod. viruses; Retroviridae; Betaretrovirus.
OX NCBI_TaxID=73490;
RN [1]
RP SEQUENCE FROM N.A.

```

Db	441	LAEVVLNRRGLDFTAEÖGGICLALAEKCCFVANKSGIVRDKIKALQEDLEKRREKTIID	500
Qy	435	TGPGWLLSOWMPWILPFLGPAATIIILLLEFGPICFNLVNFVSSIEAVK	484
	1		
	1		
	1		
Db	501	NPFWTGHGHLPGYLLPLRLPLRCLLLITTEPFLFNMTIAVKKQMAIQ	550
RESULT 4			
ENV_SRV1			
ID	ENV_SRV1	STANDARD;	PRT; 587 AA.
AC	P04027;		
Dt	23-OCT-1986 (Rel. 02, Created)		
Dt	23-OCT-1986 (Rel. 02, Last sequence update)		
Dt	16-OCT-2001 (Rel. 40, Last annotation update)		
DE	ENV polypotein precursor (Coat polypotein) [Contains: Coat protein GP20].		
DE	GP20; Coat protein GP20].		
GN	ENV.		
OS	Simian retrovirus SRV-1.		
OC	Viruses; Retroid viruses; Retroviridae; Betaretrovirus.		
OX	NCBI_TaxID=11942;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=86151668; PubMed=3006247;		
RX	Power M.D., Marx P.A., Bryant M.L., Gardner M.B., Barr P.J.,		
RA	Lucini P.A.;		
RT	"Nucleotide sequence of SRV-1, a type D simian acquired immune		
RT	deficiency syndrome retrovirus.";		
RL	Science 231:1567-1572(1986).		
CC	-1-PMW:SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use, by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/		
CC	or send an email to license@isb-sib.ch).		
CC	-----		

DR	HSSP; P03385; IMOF.	Env-Polyprotein.
DR	InterPro; IPR002050;	Env-Polyprotein.
DR	Pfam; PF00429; ENV_Polyprotein; 1.	
KW	Coat protein; Glycoprotein; Polyporein; Transmembrane; Signal.	
FT	SIGNAL	1 20
FT	CHAIN	21 395
FT	CHAIN	396 587
FT	TRANSMEM	399 419
FT	TRANSMEM	528 548
FT	CARBOHYD	120 120
FT	CARBOHYD	237 237
FT	CARBOHYD	266 266
FT	CARBOHYD	271 271
FT	CARBOHYD	277 277
FT	CARBOHYD	280 280
FT	CARBOHYD	295 295
FT	CARBOHYD	308 308
FT	CARBOHYD	322 322
FT	CARBOHYD	328 328
FT	CARBOHYD	340 340
FT	CARBOHYD	358 358
FT	CARBOHYD	488 488
FT	CARBOHYD	585 585
SO	SEQUENCE	587 AA; 64474 MW; 80933DD3BF85ABD CRC64;

Query Match	20.9%; Score 599; DB 1; Length 587;
Best Local Similarity	30.2%; Pred. No. 2e-39;
Matches 164; Conservative %	86; Mismatches 191; Indels 102; Gaps 20;
24	PCRC---MTSSSPQGEFLMRQGRGNDAPYSRLS-----KCPPTTATH-----67
DB	41 PCDCAGGVSSPPNLSLT-----TVSCSTYTAIVSYTNLSLKWCVCSTPTTASPTHTGSC 93
DP	-----MPNCHSATLCHNANTHYWTGMI--NPS-----95

```

Db      94  PSCNSQSYDSVHATCYNHNQOCTIGNKTYLFTATMIDKSPSSDGNVPTLGNQMLIT 153
QY      96  --CP--GGLGVYTCMTYTTOTGMSDGGVQDQAREKHKEVYSQLTRVHGHSPPKGLDLS 152
Db      154  AGCPENKKGVCWNSQPSVHMSDGGPQDKVREIYNNKFEELHKLFPPELSIHPDALP 213
QY      153  KLN--ETLRHTRLVLSFN--TTLTGHEVSAQN--PTNWCICLPNFRPVYSIPVPEOW 206
Db      214  EARGKEKIDANT-----FDLLATVHSLNLSQROLAEDDWCL-----RSGDPVPLALPY 264
QY      207  NMFSTEINT-----TSVLGP--LVSNLEITHSNLTCV--KESNTYV-----T 247
Db      265  DNTSCNSTEFFNCSNCSCLITPPFLVQPNFTHS---VCLADYQNNSPDIDVGLAGFT 321
QY      248  NSQICRNVPTPTQVCLPSCIFPYCNS--AYRCNLSSESMCLFLVPMITTYTED-- 304
Db      322  NCSSYINISKPSSEPLCAPNSSVFVCGNNKAYTLPTMTGSCVLAITLLPIDITIPGSEPV 381
QY      305  ---LYSVISKPRNKRVPILPFVIGAGVLGALGTGIGITTSQFYKLSQELNGDERV 361
Db      382  PIAIDHFLGRPK--RAIQFIPVIGLITTAVSTGTAGLSVLQYTKLSHQILSDYQAI 440
QY      362  ADSLVTLQDQNLAAVYLQNRALDLTAREGSTCLFLICECCYVYNSGITYEKYEI 421
Db      441  SSTIQDQDQVDSLAEVYLQNRKRLDLTAEQGCICLAEKCCFYANKSGIYADKIKNL 500
QY      422  RDRIORAEELRTMGPGGLTSQWMPILPFLGPLAATILLFPGCIFLNVVSSRIE 481
Db      501  QDLEKRRKQIDINPFTGHEGLLPYMPILGLLCLLVLSGPIITFNKLMFTIKQIE 560
QY      482  AVK 484
Db      561  SIQ 563

```

RESULT 5

ENV_AVISN STANDARD; PRT; 567 AA.

```

AC P31796:
DT 01-JUL-1993 (Rel. 26, Last Created)
DT 02-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE ENV polyprotein (Coat polyprotein) [Contains: Coat protein GP73; Coat
  protein GP22].
GN ENV.
OS Avian spleen necrosis virus.
OC Viruses; Retrofold viruses; Retroviridae; Avian type C retroviruses.
OX NCBI_TaxID=11899;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Clone PB101;
RX MEDLINE=92219390; PubMed=1313915;
RA Kewalramani V.N., Pangniban A.T., Emerman M.;
RT "Spleen necrosis virus, an avian immunosuppressive retrovirus, shares
  a receptor with the type D simian retroviruses."
RL J. Virol. 66:3026-3031(1992).
CC -1- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL Outstation
  at the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC
CC EMBL; M87666; ; NOT_ANNOTATED_CDS.
DR PIR; A38212; VCFVAS.
DR HSSP; P03385; IMOF.
DR Interpro; IPR002050; Env_Polyprotein.
DR Pfam; PF00429; Env_Polyprotein; 1.
KW Coat protein; Glycoprotein; Transmembrane; Polyprotein.

```

```

FT CHAIN 1 397 COAT PROTEIN GP73.
FT CHAIN 398 567 COAT PROTEIN GP22.
FT TRANSMEM 402 418 POTENTIAL.
FT CARBOHYD 245 245 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 274 274 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 280 280 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 306 306 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 319 319 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 328 328 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 335 335 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 491 491 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 567 AA; 61596 MW; 204EA57C32159175 CRC64;

```

Query Match

Best Local Similarity 20.7%; Score 593.5; DB 1; Length 567;
Matches 154; Conservative 75; Mismatches 176; Indels 55; Gaps 13;

```

QY 66 THMRNCHSATLGMANTHYWT-----GKMINPSCPGGLGVYTCWTFY 110
Db 121 TOMHSTCYEKTQECTLLGKTYFTALQKTKLGSYEDGPNKLIQASCTGTGKPCVQMPVA 180
QY 111 QTGMSDGGVQDQAREKHKEVYSQLTRVHGTSPPYGLDLSKLEHTRTLVSLFNT 170
Db 181 PVYVSDGGPDMIRRESVRELEIIRHSYPSVQYHPALPR-----SKGVLDLPQ 232
QY 171 T--LTGLHEV-SAQN--TNCWICLPNFRPVYSIPVPEOWNMFSTEINTTSVLVGPL 223
Db 233 TSDILEATHQVLNATNFKLNMCLMCLCTCTPIPALP-----TNCVATLDGNCISLPLF 287
QY 224 VSN---LEIT---HTSNLTCKVFSNTYTTNSQICRWATPPQ-----YCLPSGIF 269
Db 288 GCPNPGSIDVSCYAGEADNRTGIPVGVHF--TNTSIOEVTNETSOMGNLRLCPPGHV 346
QY 270 FVCGTS--AYRCNLSSESMCLFLVPMITTYTED--LYSVISKPRNKR--VPILPV 324
Db 347 FVCGNNMAYTALPKKATGICLILASIVPDIISISEEPIPLPSLEYTRARRKRAVOFTPL 406
QY 325 IGAGVLGALGTGIGITTSQFYKLSQELNGDERVADSLVTLQDQNLAAVYLQNR 384
Db 407 VGLGISGATLAGGGLGVSVHTYHKLSNGLIEDVQALSGTINDLQDIDSLAEVYLQNR 466
QY 385 ALDLTAREGSTCLFLICECCYVYNSGITYEKYERDRIOARRAEELRTMGPGGLSOW 444
Db 467 GLDLTAREQGCICLAEKCCFYANKSGIYADKIRKIQEDLIEKKRALYDNPMSGLNGF 526
QY 445 MPMLPFLGPLAATILLFPGCIFLNVVSSRIEAVK 484
Db 527 LPYLLPLGLPFLGLILFLTLGPCIMKTTLRIINDKIDAVK 566

```

RESULT 6

ENV_MPMV STANDARD; PRT; 586 AA.

```

AC P07575;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE ENV polyprotein precursor (Coat polyprotein) [Contains: Coat protein
  GP70; Coat protein GP20].
GN ENV.
OS Simian Mason-Pfizer virus (MPMV).
OC Viruses; Retrofold viruses; Retroviridae; Betaretrovirus.
OX NCBI_TaxID=11855;
RN [1]
RP SEQUENCE FROM N.A. (CLONE 6A).
RX MEDLINE=8618951; PubMed=2421920;
RA Sonligo P., Barker C., Hunter E., Wain-Hobson S.;
RT "Nucleotide sequence of Mason-Pfizer monkey virus: an
  immunosuppressive D-type retrovirus."
RL Cell 45:375-385(1986)
CC -1- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```


CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

CC -----
 CC EMBL: M12349; AAA47712.1; -
 CC PIR: D25839; VCLJMP.
 CC HSSP: P03385; 1MOP.
 DR InterPro: IPR002050; Env_polyprotein.
 DR Pfam: PF00429; Env_polyprotein; 1.
 DR Coats: protein; Glycoprotein; Polyprotein; Signal.
 FT SIGNAL 1 20
 FT CHAIN 21 394 COAT PROTEIN GP20.
 FT CARBOHYD 120 586 COAT PROTEIN GP20.
 FT CARBOHYD 120 120 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 237 237 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 264 264 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 276 276 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 291 291 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 304 304 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 318 318 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 324 324 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 357 357 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 487 487 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 586 AA; 63882 MW; 3CE7A399D9E2450 CRC64;

Query Match 20.7%; Score 593.5; DB 1; Length 586;
 Best Local Similarity 28.6%; Pred. No. 5.5e-39;
 Matches 156; Conservative 89; Mismatches 192; Indels 109; Gaps 16;

OY 24 PCRMTSSSPYOFPLMRQRPNDAPSRKSK-----GTPPT 63
 DB 41 PCDC-----AGGYSSPTINSLTVCSTHTAYSVTSLKQCVSTPTTP 85
 OY 64 AHTH-----MPNCHSATLCKMANTHWYTG----- 89
 DB 86 SNTHTGSCRGECNTSYDSYHASCYNHOQCNGIKKTYTLATITGDRPTAIDGAVPYLV 145
 OY 90 ---KMINSPCG-LGYVTCWTFYTOGMSDGGVODAREKHVEKVISOUTLRVHGTS 144
 DB 146 GTHMLITAGCNGKKGVCWNSRPSVHISDGGGPODKARDIYKKEELHRSLEPEL 205
 OY 145 PYKGLDLSKLN--ELRTHF-----RLVSLFNTITLGLHEVSAQNPTNCWICLPNFRY 197
 DB 206 STHPLALPEARKKEKIDATHTDLATVHSLNLAOSPLAE-----DCMLCIGSQDPVP 258
 OY 198 VSIPPEO-WNNFSTEINTSVLWGP-LVSNLEIHTSNTLVCVKSNTTYTN-----S 249
 DB 259 LALPYNDILCSNFALSHNSCLTPPLVQPNFT--DSCLLAHYQNNSEFDIDVGLAST 317
 OY 250 QC-----IRWTPPQIVCLPSGIFVCGTS--AYRCLNGSSSESMCFSLVPMPTIYEQ 303
 DB 318 NCSYYNTASTKSPNSLCAPNSYVCGNNKRAYTYLPTNMGCVLATLLDPDIDIPS 377
 OY 304 D-----LVSYSKRNRKRVPLPFVIGAGVLGALCTGGTTSQTFYRKISQELNGDM 358
 DB 378 EEPVPIPAIDHFLGKAK-RAIQIPLFVGGITTAVSTGAAGVSTYOTYTKLSHDLISV 436
 OY 359 EKVADSLVLDOLSLAAVYLVONRRALDLTAERGICLFLGEECCYVNGSVTEKX 418
 DB 437 QAISSITIDLDQVDSIAEVYLVONRGDILTAEGGICLALQEKCCFFANSGIYRDXI 496
 OY 419 KEIRRIORRAEELNATNGWGLLSQMMWILPFLGPLAAIILLLFGPCIFLNVYSS 478
 DB 497 KNLQDLERRRQLIDNPFMTSFGFLPYVMDILGPLICLLIYLSFGPIIFKMTFIKH 556
 OY 479 RIEAVK 484
 DB 557 QIESIQ 562

RESULT 7
 ENV_SMRVH
 ID ENV_SMRVH STANDARD; PRT; 575 AA.
 AC P21412;
 DT 01-MAY-1991 (Rel. 18, Created)
 DT 01-MAY-1991 (Rel. 18, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE ENV polyprotein precursor (Coat polyprotein) [contains: Outer membrane
 DE protein; Coat protein GP20].
 GN ENV.
 OS Squirrel monkey retrovirus (SMRV-H) (SMRV-HLB).
 OC Viruses; Retroid viruses; Retroviridae; Betaretrovirus.
 OX NCBI_Taxid-11856;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-89073750; PubMed-3201749;
 RA Oda T., Ikeda S., Watanabe S., Hatsushika M., Akiyama K.,
 RA Matsunobu F.;
 RT "Molecular cloning, complete nucleotide sequence, and gene structure
 RT of the provirus genome of a retrovirus produced in a human
 RT lymphoblastoid cell line.";
 RL Virology 167:468-476(1988).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

CC EMBL: M23385; AAA66454.1; -
 CC EMBL: M23385; AAA66455.1; -
 DR PIR: D31827; VCLJHD.
 DR HSSP: P03385; 1MOP.
 DR InterPro: IPR002050; Env_polyprotein.
 DR Pfam: PF00429; Env_polyprotein; 1.
 DR Coats: protein; Glycoprotein; Polyprotein; Transmembrane; Signal.
 FT SIGNAL 1 17
 FT CHAIN 18 386 OUTER MEMBRANE PROTEIN.
 FT CHAIN 387 575 COAT PROTEIN GP20.
 FT TRANSMEM 387 403 POTENTIAL.
 FT SITE 518 534 POTENTIAL.
 FT SITE 447 481 IMMUNOSUPPRESSIVE PEPTIDE.
 FT CARBOHYD 126 126 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 239 239 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 266 266 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 271 271 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 302 302 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 316 316 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 322 322 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 349 349 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 479 479 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 575 AA; 62245 MW; 7210DECC53896669 CRC64;

Query Match 19.1%; Score 549; DB 1; Length 575;
 Best Local Similarity 30.6%; Pred. No. 1.7e-35;
 Matches 148; Conservative 83; Mismatches 148; Indels 105; Gaps 18;

OY 66 THMPNCHSATLCKMANTHWYTGK-----INP-----SCPGGLGYTV 104
 DB 109 SQMHSSCVSFPSCQOGNNTYTLALQRTKSNSENPVYSGIHPHVLQAGDGTVGKSV 168
 OY 105 CWTFTTQIGMSDGGVODAREKHVEKVISOUTLRVHGTS-----GTSSPYGLDLSKLN 155
 DB 169 CWNQOAPHVSVSGGGPODAVRELYVOKIETLYIOSQFPLSYHPLARSKP-RGPDID 224
 OY 156 ETLRTHTRLVSLFNTITLGLHEVSAQNPTNCWICL-----PLNF----- 194
 DB 225 -----AQMLDILSATHOALNINSPSLAONCHILCNGTSMPLAPVNISSFNASQNNCT 278
 OY 195 --RPVYSIPVPEO-WNNF--STEINTSVLWGPVLSNLEIHTSNTLVCVKSNTTYTNS 249

```

Db      279 PSLPFRVQPMPSQVPCFFKGAONNSFIDPVG--VAN-----FVNCSSSNH 323
OY      250 QCIMWVTPPIQIVCLPSGIFVCGTS-AYRCLNSSSEMCFLSVLPMTTYTQD---- 304
Db      324 -----SALCPGPGQAVCGNNLAFTALPANMTGSCVLAALLPIDIDISDDDPVI 374
OY      305 -LASYISKPRNKRPVLPFEVIGAGVLCALGTGIGITTSQFYKLSQELNGMERVAD 363
Db      375 PTFEYVIGR-QKRAVTLPLVIGVSTAVATGAGLVANQSTKLSHOLINDVALSS 433
OY      434 TINDLODOLSLAVLQNRRLDLTAERGCTCLFGECCYVYVNGSGIVTEKVEIRD 423
OY      424 RIORR-----ABELNTGPMGLSQMMPMILPEGLPLAIIILLFGCITFLVNFVSSR 479
Db      494 DLEKRRKALNDNLTLGLNGLL-----PYLPLFLPLRAIILFFSFAPIILRRVALLRDO 549
OY      480 IEAV 483
Db      550 LNSL 553

```

RESULT 8

ENV_AVIRE STANDARD; PRT; 582 AA.

```

AC P03399:
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE ENV polypeptide (Coat polypeptide) [Contains: Coat protein GP73; Coat
  protein GP22].
DE ENY.
OS Avian reiculoendotheliosis virus.
OC Viruses; Retrovird viruses; Retroviridae; Gammaretrovirus.
OX NCBI_TaxID=11636;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=85009850; PubMed=6090694;
RA Wilhelmsson K.C., Eggleston K., Temin H.M.;
RT Nucleic acid sequences of the oncogene v-rel in
  reiculoendotheliosis virus strain T and its cellular homolog, the
  proto-oncogene c-rel."
RT J. Virol. 52:172-182(1984).
RL -1- PPM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- MISCELLANEOUS: STRAIN A IS A HELPER VIRUS OF THE STRAIN T.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL: K02537; AAA99198.1; -
DR EMBL: X01455; CAA25686.1; -
DR PIR: A03999; VCVDAK.
DR HSP: P03385; IMOF.
DR InterPro: IPR002050; Env_Polypeptide.
DR Pfam: PF00429; Env_Polypeptide.1.
KW Coat protein; Glycoprotein; Transmembrane; Polypeptide.
FT CHAIN 1 391
FT CHAIN 392 582
FT CARBOHYD 241 301
FT CARBOHYD 301 301
FT CARBOHYD 314 314
FT CARBOHYD 485 485
FT SEQUENCE 582 AA; 6438 MW; CD2560ADF0C6D52 CRC64;

```

Query Match

17.5%; Score 502; DB 1; Length 582;

Best Local Similarity 28.6%; Pred. No. 8.7e-32;
Matches 159; Conservative 75; Mismatches 192; Indels 130; Gaps 19;

```

OY      13 LPSFTLTAPPCCMCMTSSSPYQELMRQRPNDAPSVRSLSKGPPTTAHTMPRNC 72
Db      101 IIPSPQEPQCECLTA-----MHAN-THYTGKMINPSCPGGLGVY 104
OY      73 YHSATLC-----
Db      126 YKAOECTLGLKTYFTALLOKTLGSTEDEGNKLLQASCTGIWETSMIGPRCP-----CV 180
OY      105 CWTYFTGTGMSDGGVDOAKERKHYEVISQTFVHGTSPPKGLDSKILHETLRTHTRL 164
Db      181 C-----LDGGGPDRF-GRICAGLEELIINHSVPQYHPLALPR-----PRG 222
OY      165 VSEFNT---LTGLHEV-SAONP---TNCVCLPLNRPVSYIVPEOMNPFETINTS 217
Db      223 VDLDPQTSDLLEATHQVLMATNPQALAEWCWLCMTLG---TSPQPSRRMMS-----LS 273
OY      218 VLVGPLVSNLEITHTSN---LTC-----VKFSNTYTSNQCIRWVTPPTQI 261
Db      274 MEIAVLASLSGATHRVNRCOLCREADNRGCIPIGYVHFTNCTIOESLRVRYELRD 333
OY      262 VCLPSGLFVCGTSAYRCLNNGSSSEMCFLSEIVPMTTYEQD---LYSVISKPRKR- 317
Db      334 YLHRVMTLCVEQHAATYALPKRWIGLCIASIVPDMSTIPGEPIPLPSIEYTAGRHKRA 393
OY      318 VPLPFIYAGVLCALGTGIGITTSQFYKLSQELNGMERVADSLVTLQDOLNSIAA 377
Db      394 VQFPLVGLGIGITATLGGTGLGVSVHTYKLSNOLIEDVALINDLODDISLAE 453
OY      378 VLQNRRLDLTAERGCTCLFGECCYVYVNGSGIVTEKKEIRDRIORREELRMTGP 437
Db      454 VVLQNRRLDLTAERGCTCLFGECCYVYVNGSGIVTEKKEIRDRIORREELRMTGP 513
OY      438 WGLSQMMPWILPEGLPLAIIILLFGCITFLNLFVNSRIKRAV-LQMEPRKQSKT 496
Db      514 WNGLNGFLPYLLPELGLPLFLTLGTCIRKTLRIHDKIGSKNPRISPAVQATP- 572
OY      497 IYRRPLDRPASRSDY 512
Db      573 -----NRDGYPRSMV 582

```

RESULT 9

ENV_RMCFV STANDARD; PRT; 640 AA.

```

AC P06445: Q85628; Q85629; Q89529;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE ENV polypeptide precursor (Coat polypeptide) [Contains: knob protein
  GP70; Coat protein PL2; Coat protein PL5E].
DE ENY.
OS Rauscher mink cell focus-inducing virus.
OC Viruses; Retrovird viruses; Retroviridae; Mammalian type C retroviruses.
OX NCBI_TaxID=11784;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=85237696; PubMed=4009793;
RA Vogt M., Haegbloom C., Swift S., Haas M.;
RT "Envelope gene and long terminal repeat determine the different
  biological properties of Rauscher, Friend, and Moloney mink cell
  focus-inducing viruses."
RT J. Virol. 55:184-192(1985).
RL -1- PPM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- MISCELLANEOUS: STRAIN A IS A HELPER VIRUS OF THE STRAIN T.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

or send an email to license@lsb.sib.ch.

CC EMBL: M10100: AAA46528.1: -

DR PIR: A03990: VCMV.RV.

DR HSP: P03385: 1MOF.

DR InterPro: IPR002050: Env.-polyprotein.

DR Pfam: PF00429: ENV_polyprotein.1.

KW Coat protein; Glycoprotein; Polyprotein; Signal.

FT SIGNAL 1 32

FT CHAIN 33 443 KNOB PROTEIN GP70.

FT CHAIN 444 623 COAT PROTEIN P12E.

FT CHAIN 444 623 COAT PROTEIN P15E.

FT CARBOHYD 43 43 N-LINKED (GLCNAc. . .) (POTENTIAL).

FT CARBOHYD 58 58 N-LINKED (GLCNAc. . .) (POTENTIAL).

FT CARBOHYD 300 300 N-LINKED (GLCNAc. . .) (POTENTIAL).

FT CARBOHYD 332 332 N-LINKED (GLCNAc. . .) (POTENTIAL).

FT CARBOHYD 339 339 N-LINKED (GLCNAc. . .) (POTENTIAL).

FT CARBOHYD 372 372 N-LINKED (GLCNAc. . .) (POTENTIAL).

FT CARBOHYD 408 408 N-LINKED (GLCNAc. . .) (POTENTIAL).

SQ SEQUENCE 640 AA: 70071 MW: 1E4450343643D799 CRC64:

Query Match 14.8%; Score 424; DB 1; Length 640;

Best Local Similarity 30.0%; Pred. No. 1.3e-25;

Matches 111; Conservative 63; Mismatches 144; Indels 52; Gaps 10;

160 THTRLVSLFNTLTGLHEVSAONPTNCWICLPINFRPYKIPVPEOMNFTSINTSVL 219

284 TGDRLNLVDAQVQALNLSPPKQECMCLVAEPYKGVAVLGTSMHTSAPNCSVA 343

220 VGPLVSNLEIT-----HTSNLCYKESNTTYTNSOCIRMTPTQIVCLPSGIF 269

344 SOHKLTLESEVTOGCLIGTVPKTHQALC-----NTTLKTKN-----GSYYLVAPAGTY 391

270 FVCGTSARVCLNGS-----SESMCLSLVPPMTITTEDLYSVYSKPNKRVPI---LP 322

332 WACNTGLPCLSAIYLNRTDYCVLELMPRTYTPPSTVYSQFEKSYHKKEPVSLLTA 451

333 FVIGAGVIGALGTGIGITTT---STQFYKLSQELNGDMERVADSLVTLQDQLNSLAAY 379

452 LLLGGLTNGGILAGVGTGTLVATVQGFQQLHAAYQDDLKEVEKSTINLEKLTSLSEV 511

360 LQNRALDLTLARERGCTCLFLGEECCYYVNSGIVTEKYKEIRDRIOARRAELRNTGPV 438

512 LQNRGLDLFLKEBGLCAALKECCFYADHGTGLVDSMAKREBLTGROKLFESSQGW 571

439 -GLLSQWMPW---LTPFLGPIAIIILLLGPGCLFENLLVNVSSRIEYK----- 484

572 EQLFNK-SPWFTTLISTINGPLIILLLILFLGFCILNRLVQFKDRISYVALVLTQYH 630

485 ---LQMEPK 490

631 QLKPLEYEPQ 640

RESULT 10

ENV_MLYMO STANDARD: PRT: 665 AA.

AC P03385; 21-JUL-1986 (Rel. 01, Created)

DT 21-JUL-1986 (Rel. 01, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE ENV_polyprotein precursor (coat polyprotein) [Contains: knob protein GP70; spike protein p15e; R protein].

GN ENV.

OS Moloney murine leukemia virus.

OC Viruses; Retrovirdae; Retroviridae; Gammaetrovirus.

OX NCBI_TaxID=11801;

RN MEDLINE FROM N.A. (CLONE PMLV-1).

RX SHIMNICK T.M., LERNER R.A., SUTCLIFFE J.G.;

RT "Nucleotide sequence of Moloney murine leukemia virus.";

RL Nature 293:543-548(1981).

RN SEQUENCE OF 484-665 FROM N.A. (PROVIRUS).

RX MEDLINE-810523872; PubMed-6159543;

RA Sutcliffe J.G., Shimnick T.M., Green N., Liu F.-T., Niman H.L., Lerner R.A.;

RT "Chemical synthesis of a polypeptide predicted from nucleotide sequence allows detection of a new retroviral gene product.";

RL Nature 287:801-805(1980).

RN SEQUENCE OF 484-665 FROM N.A. (PROVIRUS; CLONE PMLV-201).

RX MEDLINE-81013872; PubMed-6251454;

RA Sutcliffe J.G., Shimnick T.M., Verma I.M., Lerner R.A.;

RT "Nucleotide sequence of Moloney leukemia virus: 3' end reveals details of replications, analogy to bacterial transposons, and an unexpected gene.";

RL Proc. Natl. Acad. Sci. U.S.A. 77:3302-3306(1980).

RN SEQUENCE OF 470-489 AND 598-665.

RX MEDLINE-82082389; PubMed-6947213;

RA Green N., Shimnick T.M., Witte O., Ponticelli A., Sutcliffe J.G., Lerner R.A.;

RT "Sequence-specific antibodies show that maturation of Moloney leukemia virus envelope polyprotein involves removal of a COOH-terminal peptide.";

RL Proc. Natl. Acad. Sci. U.S.A. 78:6023-6027(1981).

RN X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS) OF 514-567.

RX MEDLINE-96196884; PubMed-8612078;

RA Fass D., Harrison S.C., Kim P.S.;

RT "Retrovirus envelope domain at 1.7-A resolution.";

RL Nat. Struct. Biol. 3:465-469(1996).

CC -1- PTM: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.lsb.sib.ch/announce/> or send an email to license@lsb.sib.ch).

CC EMBL: J02255; AAB5943.1: -

DR PIR: A93265; VCMVEM.

DR PDB: 1MOF: 14-OCT-96.

DR InterPro: IPR002050: Env.-polyprotein.

DR Pfam: PF00429: ENV_polyprotein.1.

KW Glycoprotein; Coat protein; Polyprotein; Transmembrane; Signal; 3D-structure.

KW SIGNAL 1 33

FT CHAIN 34 469

FT CHAIN 470 649

FT CHAIN 650 665

FT CHAIN 650 665

FT TRANSMEM 610 627

FT DISULFID 555 562

FT CARBOHYD 45 45 N-LINKED (GLCNAc. . .) (POTENTIAL).

FT CARBOHYD 199 199 N-LINKED (GLCNAc. . .) (POTENTIAL).

FT CARBOHYD 326 326 N-LINKED (GLCNAc. . .) (POTENTIAL).

FT CARBOHYD 358 358 N-LINKED (GLCNAc. . .) (POTENTIAL).

FT CARBOHYD 365 365 N-LINKED (GLCNAc. . .) (POTENTIAL).

FT CARBOHYD 398 398 N-LINKED (GLCNAc. . .) (POTENTIAL).

FT CARBOHYD 434 434 N-LINKED (GLCNAc. . .) (POTENTIAL).

FT CONFLICT 655 655 Y -> F (IN REF. 3).

FT CONFLICT 663 663 Y -> C (IN REF. 3 AND 4).

FT HELIX 516 547

FT TURN 548 548

FT HELIX 549 551

FT TURN 552 552

FT HELIX 554 558

FT TURN 559 559

SQ SEQUENCE 665 AA: 73301 MW: 12EBA09C8FB93FE2 CRC64:

Query Match 14.6%; Score 418.5; DB 1; Length 665;

Best Local Similarity 30.2%; Pred. No. 3.8e-25;

FT	SEQUENCE	662 AA;	73149;MM;	1482080547CFE47	CRC64;	TS65VAQARQROQTNSAARSAVPTMG (IN REF. 4).
SQ	SEQUENCE	662 AA;	73149;MM;	1482080547CFE47	CRC64;	TS65VAQARQROQTNSAARSAVPTMG (IN REF. 4).
	Query Match	14.4%;	Score 412.5;	DB 1;	Length 662;	
	Best Local Similarity	23.2%;	Pred. No. 1,1e-24;			
	Matches 136;	Conservative 89;	Mismatches 208;	Indels 161;	Gaps 21;	
QY	44 PG-NIDAPYSRLSGTPPTFAHTM-----PRNCYHSATLCMHANTHYMGKMINPS	95				
DB	94 PGYGCDDQPMRMWQQRNPPFYVCPCPHANRKKQCGGPDGFCAMVGGCTTGTYW-----RPT	148				
QY	96 -----CPGGAGVTCW---YFTGTGMSDGG-----	119				
DB	149 SSMWDITVKKGVGTGIGYQSGGCGPCYDKAVHSSPTTGASEGRCPNLILOFTOKROT	208				
QY	120 -----VQD-----AREKHKEVI	133				
DB	209 SMDGPKSMGLRLYNSGDPLALEFSVSNQWHTIPRQAMGRVLVLPDOKPPRSQOIESRV	268				
QY	134 SQ-----LTVHGTSSPYKGLDLSKMETLRTNRLVSLFNLTTLTGHEVSAQN	182				
DB	269 TPHHSQNGGTGPGITLVNASIAPLSTPTVPTRASPFRIGTGRILNLVQGTVALNATDPNR	328				
QY	183 PTNCWICLLPLNFRYVSLPPEQNNNSTELNTSVLVGRLVSLDELTHHSNLTGVAFSN	242				
DB	329 TKDCWCLVLSRPPEYEGIAL--LGNYSNQTN-----PPPSCLSTPQH-----KLTI	372				
QY	243 TTYVTNSQCRIMVYPPRQIVC-----LPSGIFCYGTSAYRC-----LNGS	283				
DB	373 SEVSGGGLCIGTVPKTHQALCNENYQOQHTGHAHYLAANGTYMACNTGTRISMAVLNMT	432				
QY	284 SESNCFSLFVLPMTTITTEQDLYSVYSIKPRNKRVP--LPEVIG--AGVLGALGTG	336				
DB	433 SD-FCVILTELPRTVYHQPEVYVYHFAKAFRRREPISLTFVALMLGGLTVGGIAAGVTG	491				
QY	337 IGGTTSTQRYFKLSQELNGDMENVASLTVLQDQLNSLAVALVQNRAALDLTAERGT	396				
DB	492 TKALIEFAQF-ROIQAMAMHTDIOALESSISALESSLSLSEVYVQNRKGDITFLQEGGL	550				
QY	450 PFLGPLAAILLLLEFGCIFNLVNFVSSRIEAVKLOMPERAKSKTRYYRRDLRP	505				
DB	609 SIMGPLILLILLILFEGCILNRLVQFYKDRISVQALILNQYQIQY--DPRP	662				
RESULT 13						
ENV_MLVF5	STANDARD;	PRT;	675 AA.			
AC	P03390;					
DT	21-JUL-1986 (Rel. 01, Created)					
DT	01-AUG-1992 (Rel. 23, Last sequence update)					
DT	28-FEB-2003 (Rel. 41, Last annotation update)					
DE	ENV polypeptide precursor [Contains: knob protein GP70; Spike protein					
DE	P15E; R protein].					
GN	ENV.					
OS	Friend murine leukemia virus (isolate 57) (F-MuLV).					
OC	Viruses; Retrod viruses; Retroviridae; Mammalian type C retroviruses.					
OX	NCBI_Taxid=11796;					
RN	SEQUENCE FROM N.A.					
RA	Zimmerman W., Koch W., von Maydell-Livonius U., Schrewe H.,					
RL	submitted (xxx-1991) to the EMBL/GenBank/DBJ databases.					
RN	[2]					
RP	PRELIMINARY SEQUENCE FROM N.A.					
RX	MEDLINE=84138778; PubMed=6321768;					
RA	Koch W., Zimmermann W., Olliff A., Friedlich R.W.,					
RT	"Molecular analysis of the envelope gene and long terminal repeat of					
RT	Friend mink cell focus-inducing virus: implications for the functions					
RT	of these sequences.";					

RL J. Virol. 49:828-840(1984).
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 43-270.
 RX MEDLINE-97433384; PubMed-9287219;
 RA Paas D., Davey R.A., Hamson C.A., Kim P.S., Cunningham J.M.,
 RA Berger J.M.;
 RT "Structure of a murine leukemia virus receptor-binding glycoprotein
 RT at 2.0-A resolution."
 RL Science 277:1662-1666(1997).
 CC -1- FUNCTION: INVOLVED IN HOST RANGE; VIRUS RECEPTOR BINDING; CELL
 CC FUSION.
 CC -1- PPM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD NATURE PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X02794; CAA26561.1; -
 DR PDB: 1AOL; 15-OCT-97.
 DR InterPro: IPR02050; Env_Polyprotein.
 DR Pfam: PF00429; Env_Polyprotein; 1.
 KW Coat protein; Glycoprotein; Polyprotein; Transmembrane; Signal;
 KW 3D-structure; Zinc; Metal-binding.
 FT SIGNAL 1 34
 FT CHAIN 35 479 BY SIMILARITY.
 FT CHAIN 480 659 KNOB PROTEIN GP70.
 FT CHAIN 660 675 SPIKE PROTEIN P15E.
 FT TRANSMEM 620 636 R PROTEIN.
 FT METAL 55 55 POTENTIAL.
 FT METAL 80 132 ZINC.
 FT DISULFID 106 121
 FT DISULFID 107 117
 FT DISULFID 155 175
 FT DISULFID 167 180
 FT DISULFID 212 218
 FT DISULFID 346 349
 FT DISULFID 395 407
 FT DISULFID 376 430
 FT DISULFID 437 450
 FT CARBOHYD 46 46
 FT CARBOHYD 202 202
 FT CARBOHYD 336 336
 FT CARBOHYD 368 368
 FT CARBOHYD 375 375
 FT CARBOHYD 408 408
 FT CARBOHYD 444 444
 FT STRAND 44 52
 FT STRAND 54 55
 FT STRAND 58 67
 FT TURN 68 69
 FT STRAND 75 75
 FT STRAND 78 78
 FT TURN 83 83
 FT TURN 84 85
 FT HELIX 88 90
 FT TURN 93 94
 FT TURN 97 98
 FT STRAND 106 106
 FT TURN 107 108
 FT TURN 110 111
 FT TURN 115 116
 FT STRAND 118 118
 FT HELIX 119 122
 FT TURN 126 127
 FT HELIX 135 146
 FT TURN 148 149
 FT STRAND 152 156
 FT TURN 159 159
 FT TURN 162 163

FT HELIX 164 167
 FT HELIX 170 172
 FT TURN 173 174
 FT TURN 178 179
 FT STRAND 183 185
 FT TURN 187 188
 FT STRAND 197 201
 FT HELIX 205 213
 FT TURN 214 215
 FT STRAND 220 225
 FT HELIX 227 231
 FT HELIX 234 236
 FT TURN 237 237
 FT STRAND 239 245
 FT TURN 248 250
 FT STRAND 253 263
 FT HELIX 266 268
 SQ SEQUENCE 675 AA; 74024 MW; A097038E422B3D3 CRC64;
 Query Match 14.4%; Score 412; DR 1; Length 675;
 Best Local Similarity 29.8%; Pred. No. 1.3e-24;
 Matches 11; Conservative 57; Mismatches 145; Indels 60; Gaps 11;
 QY 160 TTRRLVSLFNTTLTGLHEVSAQNPTNCICLPLNFRPVYSIPYEDQNNFSTEINTSVL 219
 Db 320 TGDRLNLNVGAYVALNLTPDKTQECMLCLVSGPPYEGVAV-----LGTYSNHTSA- 372
 QY 220 VGPLVSLLETHSNLTCVKEFSNTYTTNSOCIRMTPTPOIYC-----LP 265
 Db 373 -----PANCASVASQKRLT-----SEVYGRGICIGTVPKTHQALCNTLTKIDKGSYLVAP 423
 QY 266 SGIFVCGTSAYRCLNGS-----SEMGFLSPVLPMTIYEDQYSVIAKPRKRVPI- 320
 Db 424 TGTMACNTGTLPELSATVNLRTTDYCVLVEIMPRVTHPSVYISQPEKSYRKKEPV 483
 QY 321 --LPFVIGAGVLCALGTGIGITTT--STQFYKLSQELNGDMERVADSLVTLQDQNSL 375
 Db 484 LTLALIGLTMGSIAGVGTALVATQFOQLHAAVODDLKEVEKSTINLEKSLTSL 543
 QY 376 AAVVLQRRALDLTARGGTCFLGEECCYVNVQSVIVKEKKEIDRQRAEELRNT 435
 Db 544 SEVVLQRRKLDLLEKEGGLCAALKECCFVADHTGLVDSNAKLEKRLTQOKRESS 603
 QY 436 GPW--GLISQWMPW---ILPFLCPLAAIILLFGFCINLVNFVSSRIEAVK----- 484
 Db 604 QGMFEGFLNR-SPEFTLLISTINGPLIILLIILFGCINLRVQFVKDHSVYQALVLT 662
 QY 485 -----IQMEP 489
 Db 663 QQYHQLKPLEYEP 675
 RESULT 14
 ENV_FLVLB STANDARD; PRT; 662 AA.
 ID ENV_FLVLB AC P11261; Q85515; Q85516; Q85517;
 DT 01-JUL-1989 (Ref. 11, Created)
 DT 01-JUL-1989 (Ref. 11, Last sequence update)
 DT 16-OCT-2001 (Ref. 40, Last annotation update)
 DE ENV polyprotein precursor (Coat polyprotein) [Contains: Knob protein
 DE GP70; Spike protein P15E (fragment)].
 GN ENV.
 OS Feline leukemia virus (strain B/Jambda-B1).
 OS Viruses; Retroid viruses; Retroviridae; Gammaretrovirus.
 OX NCBI_TaxID=103916;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-88036192; PubMed-2444714;
 RA Nicolaissen-Strouse K., Kumar H.P.M., Pitting T., Grant C.K.,
 RA Elder J.H.;
 RT "Natural feline leukemia virus variant escapes neutralization by a
 RT monoclonal antibody via an amino acid change outside the antibody-
 RT binding epitope."

RL J. Virol. 61:3410-3415(1987).
 CC -1- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 DR EMBL: J03448; AAA3048.1; -
 DR PIR: A27172; VCMVLB.
 DR HSSP: P03385; IMOF.
 DR InterPro: IPR002050; Env_polyprotein.
 DR Pfam: PF00429; Env_polyprotein: 1.
 DR Coar protein: Glycoprotein; Polypeptide; Signal.
 FT SIGNAL 1 33
 FT CHAIN 34 465
 FT CAROHD 466 662 SPIKE PROTEIN GP70.
 FT CAROHD 43 43 N-LINKED (GLCNAc. . .) (POTENTIAL).
 FT CAROHD 58 58 N-LINKED (GLCNAc. . .) (POTENTIAL).
 FT CAROHD 286 286 N-LINKED (GLCNAc. . .) (POTENTIAL).
 FT CAROHD 322 322 N-LINKED (GLCNAc. . .) (POTENTIAL).
 FT CAROHD 337 337 N-LINKED (GLCNAc. . .) (POTENTIAL).
 FT CAROHD 351 351 N-LINKED (GLCNAc. . .) (POTENTIAL).
 FT CAROHD 354 354 N-LINKED (GLCNAc. . .) (POTENTIAL).
 FT CAROHD 430 430 N-LINKED (GLCNAc. . .) (POTENTIAL).
 FT NON_TER 662 662
 SQ SEQUENCE 662 AA; 73132 MW; DAAC18E197230575 CRC64;
 Query Match 14.3%; Score 411.5; DB 1; Length 662;
 Best Local Similarity 27.8%; Pred. No. 1.4e-24;
 Matches 110; Conservative 69; Mismatches 164; Indels 53; Gaps 11;
 QY 136 LTRVHGTSPPKGLDLSKILHETLRHRLVSLFNTTGLHEVSAONPTNCICPLNFR 195
 DB 282 ILLVNASIAPLSTPVPAPSKRIGTGNRLINLVGTYALANTNPKTKDCWLCVSRPP 341
 QY 196 PYVSIPEPQMNNEFTSTINTSVLGVPLVSNLEHTHTSLNLCVKRSNTTYTNSCIRKY 255
 DB 342 YIEGIAV--LGNISNQTNPSPSCISDQHKLTISEVSG-----QSGSIGTV 385
 QY 256 TPPTQIVC-----LPSGIFVCGTSAVRC-----LNGSSSMCFSLVLP 296
 DB 386 PTHQALCKTKGKGGKHTHYLAAPSGTYMACNTGLPCISMAVLWMTSD-FCVLLLEWNR 444
 QY 297 MTITTEQDLSTVLSKPRNKRPV--LPEVIG-----AGVLCALGTGIGITTSQFYRK 349
 DB 445 VTYHQPEVYVTHFDKTVLRREPISLFTVALMLGVLGVIAGVGTFKALIFEAQF-CQ 503
 QY 350 LSGELNGMERVADSLVTLQDQLSLAVALVONRBAIDLTLTHERGTCLFLGECCYYN 409
 DB 504 LOMAHHTIDALEESISALEKSLTSLSEVVLONRGDILILFQEGSLCAALKECCFYAD 563
 QY 410 QSGIVTEKVEIRIDIRRAELRNTGFWGLSDWM--PW-----ILPLGFLAAIILL 462
 DB 564 HTGLVROMAKRIKRELKROQLFDSQOGM--FEGFNKSPMTTLLISSMGLIILLITL 621
 QY 463 LFGPCIFNLVNFVSSRIEAVKLOMEPRKQSKTKLY 498
 DB 622 LFGPCILNRLVQFVKDRISVQALITLQOYQOIKQY 657
 RESULT 15
 ENV_MLVLP STANDARD: . PRT: 676 AA.
 AC P26803;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE ENV polyprotein precursor [Contains: knob protein GP70; spike protein
 P15e; R protein].

GN ENV.
 OS Friend murine leukemia virus (isolate PVC-211) (F-MuLV).
 OC Viruses; Retroid viruses; Retroviridae; Mammalian type C retroviruses.
 OX NCBI_Taxid=11798;
 RN
 RP SEQUENCE FROM N.A.
 RX MEDLINE-92219364; PubMed-1560524;
 RA Masuda M., Remington M.P., Hoffman P.M., Ruscetti S.K.;
 RT "Molecular characterization of a neuropathogenic and
 RT nonerythroleukemogenic variant of Friend murine leukemia virus
 RT PVC-211".
 RL J. Virol. 66:2798-2806(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-92319660; PubMed-1620621;
 RA Remington M.P., Hoffman P.M., Ruscetti S.K., Masuda M.;
 RT "Complete nucleotide sequence of a neuropathogenic variant of Friend
 RT murine leukemia virus PVC-211".
 RL Nucleic Acids Res. 20:3249-3249(1992).
 CC -1- FUNCTION: INVOLVED IN HOST RANGE; VIRUS RECEPTOR BINDING; CELL
 CC FUSION.
 CC -1- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
 CC -1- MISCELLANEOUS: GP70 APPEARS TO BE RESPONSIBLE FOR THE
 CC NEUROPATHOGENICITY OF PVC-211 MuLV.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 DR EMBL: M93134; AAA6478.1; -
 DR PIR: A38210; VCMVPV.
 DR HSSP: P03390; IAOI.
 DR InterPro: IPR002050; Env_polyprotein.
 DR Pfam: PF00429; Env_polyprotein: 1.
 DR Coar protein: Glycoprotein; Polypeptide; Transmembrane; Signal.
 FT SIGNAL 1 34
 FT CHAIN 35 479
 FT CAROHD 480 659 SPIKE PROTEIN GP70.
 FT CHAIN 660 676 R PROTEIN.
 FT TRANSMEM 620 636 POTENTIAL.
 FT DISULFID 80. 132 BY SIMILARITY.
 FT DISULFID 106 121 BY SIMILARITY.
 FT DISULFID 107 117 BY SIMILARITY.
 FT DISULFID 155 175 BY SIMILARITY.
 FT DISULFID 167 180 BY SIMILARITY.
 FT DISULFID 212 218 BY SIMILARITY.
 FT DISULFID 346 349 BY SIMILARITY.
 FT DISULFID 395 407 BY SIMILARITY.
 FT DISULFID 376 430 BY SIMILARITY.
 FT DISULFID 437 450 BY SIMILARITY.
 FT CAROHD 46 46 N-LINKED (GLCNAc. . .) (POTENTIAL).
 FT CAROHD 202 202 N-LINKED (GLCNAc. . .) (POTENTIAL).
 FT CAROHD 336 336 N-LINKED (GLCNAc. . .) (POTENTIAL).
 FT CAROHD 368 368 N-LINKED (GLCNAc. . .) (POTENTIAL).
 FT CAROHD 375 375 N-LINKED (GLCNAc. . .) (POTENTIAL).
 FT CAROHD 408 408 N-LINKED (GLCNAc. . .) (POTENTIAL).
 FT CAROHD 444 444 N-LINKED (GLCNAc. . .) (POTENTIAL).
 SQ SEQUENCE 676 AA; 73945 MW; 18BD8A7FED0A8FA4 CRC64;
 Query Match 14.3%; Score 410; DB 1; Length 676;
 Best Local Similarity 29.2%; Pred. No. 1.8e-24;
 Matches 111; Conservative 58; Mismatches 139; Indels 72; Gaps 12;
 QY 160 TTRLVSLFNTTLTGLHEVSAONPTNCICPLNPEPVYSIPVPEQMNNEFTSTV 219
 DB 320 TQDRLLNLVQGVQALNLTNPKTKDCWLCV-SAPRYE----- 358
 QY 220 VGPLVSNLEHTHTS-NTLCV-----KFSNTTYTNSCIRWTPPTQIVC----- 263

Db 359 -GVAVLGTYSNHTSAPANCSSQHKLTLSVETGGLGIGTVPKTHOALCNTTLTKGKS 417
QY 264 ----LPSGIFVCGTSAVRCILNGS----SSMCFLSFLVPPMTIYTEODLYSYISKPRN 315
Db 418 YLYVAPAGTMACNTGLTPCLSATVLRRTDYCVLWLPRTYHPPSYYSQPEKSYRH 477
QY 316 KRVPPI---LPEVIGAGVIGALGTGIGGITT---STQFYKLSQELNGDMERADSLVTLQ 369
Db 478 KREPVSLTLALLGGLMGGLAAGVGTTALVATQOQFQULHAADVODLKEVEKSTINLE 537
QY 370 DQINSLAAVYLQNRALDILLTAERGTCLELGECCYYVNOGSGIVTEKKEIRDRIQRA 429
Db 538 KSLTSLSEVYLQNRGLDLFLKEGGLCAALKECCFYADHTGLVDRDSMAKLRERLTORQ 597
QY 430 EELRNTGPM--GLISQMPM----ILPELGPLAAIIIIILFGPCIFNLVNFVSSRIEAV 483
Db 598 KLESSQGWFEGLFNR-SPMFTTILISTIMGPLIILLILFLGPCIILNRLVOFVDRISVY 656
QY 484 K-----LQMEPK 490
Db 657 QALVLTQQYHQLKPLEYEPQ 676

Search completed: September 23, 2003, 14:55:19
Job time : 29 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 23, 2003, 14:45:57 : Search time 104 seconds
(without alignments)
1334.926 Million cell updates/sec

Title: US-10-069-883-1
Perfect score: 2871
Sequence: 1 MALPHYHFLFTVLLPSFTLT.....PPEHSAQPLLRNSAGSS 538

Scoring table: BLASTUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues
Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_protent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2871	100.0	538	4 Q9UQF0	Q9UQF0 homo sapien
2	2868	99.9	538	4 Q9NZG3	Q9NZG3 homo sapien
3	2867	99.9	538	4 Q9NRZ2	Q9NRZ2 homo sapien
4	2846	99.1	533	4 Q8NHV7	Q8NHV7 homo sapien
5	2490	86.7	542	15 Q9J1W9	Q9J1W9 multiple sc
6	2194	76.4	410	4 Q9S244	Q9S244 homo sapien
7	1663	57.9	319	4 Q96L59	Q96L59 homo sapien
8	1657	57.7	319	4 Q96L60	Q96L60 homo sapien
9	1650	57.5	319	4 Q96L61	Q96L61 homo sapien
10	1229	42.8	242	4 Q9S245	Q9S245 homo sapien
11	929	32.4	180	4 Q96T85	Q96T85 homo sapien
12	857	29.9	180	4 Q96T81	Q96T81 homo sapien
13	846	29.5	180	4 Q96T84	Q96T84 homo sapien
14	828.5	28.9	179	4 Q96T83	Q96T83 homo sapien
15	828	28.8	180	4 Q96T82	Q96T82 homo sapien
16	809.5	28.2	179	4 Q96T88	Q96T88 homo sapien

17	802	27.9	180	4 Q96T86	Q96T86 homo sapien
18	778	27.1	180	4 Q96T87	Q96T87 homo sapien
19	616.5	21.5	580	15 Q9WPZ9	Q9WPZ9 simian retro
20	612	21.3	564	15 Q98554	Q98554 rd14 retro
21	602.5	21.0	580	15 Q9WQ03	Q9WQ03 simian retr
22	586	20.4	620	6 Q9GLF7	Q9GLF7 trichosurus
23	567.5	19.8	618	11 Q8B141	Q8B141 mus musculu
24	564.5	19.7	584	4 Q9UNM3	Q9UNM3 homo sapien
25	563.5	19.6	584	15 Q9N2K0	Q9N2K0 herv-h/herv6
26	496.5	17.3	563	15 Q9N2J9	Q9N2J9 herv-h/herv6
27	488.5	17.0	584	6 Q8M1B6	Q8M1B6 pan troglod
28	484	16.9	555	15 Q9N2J8	Q9N2J8 herv-h/herv6
29	465.5	16.2	321	15 Q9S708	Q9S708 rd14 retro
30	426.5	14.9	638	15 Q901X5	Q901X5 porcine end
31	425.5	14.8	638	15 Q901X4	Q901X4 porcine end
32	425.5	14.8	641	15 Q8O6T9	Q8O6T9 porcine end
33	424	14.8	669	15 Q9YV53	Q9YV53 murine leuk
34	423.5	14.8	638	6 Q6Z705	Q6Z705 sus scrofa
35	423.5	14.8	638	15 Q901X3	Q901X3 porcine end
36	423	14.7	665	15 Q8UMZ9	Q8UMZ9 moloney mur
37	420.5	14.6	641	15 Q83363	Q83363 murine leuk
38	420	14.6	673	15 Q89816	Q89816 mus dunni e
39	419.5	14.6	637	15 Q85630	Q85630 friend mink
40	417	14.5	638	15 Q9DQ21	Q9DQ21 murine leuk
41	417	14.5	676	15 Q41251	Q41251 rauscher mu
42	416	14.5	635	15 Q9WHJ7	Q9WHJ7 friend mink
43	414.5	14.4	614	15 Q83380	Q83380 rat leukeml
44	414.5	14.4	636	15 Q85506	Q85506 murine leuk
45	413.5	14.4	455	15 Q90120	Q90120 mink cell f

ALIGNMENTS

RESULT 1	Q9UQF0	PRELIMINARY:	PRT:	538 AA.
ID	Q9UQF0			
AC	Q9UQF0			
DT	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DE	01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
DE	Envelope protein precursor.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Placenta;			
RX	MEDLINE=99099005; PubMed=9882319;			
RA	Blond J.L., Beseme F., Duret L., Bouton O., Bedin F., Perron H.,			
RA	Mandrand B., Mallet F.;			
RT	"Molecular characterization and placental expression of HERV-W, a new			
RT	human endogenous retrovirus family.";			
RL	J. Virol. 73:1175-1185(1999).			
EMBL	AF072506; AAD14546.2; -.			
DR	HSPF; P03385; 1MOF.			
DR	InterPro: IPR02050; Env_polypotein.			
DR	Pfam: PF00429; Env_polypotein. 1.			
KW	SIGNAL.			
FT	SIGNAL.			
FT	CHAIN 1 21 POTENTIAL.			
FT	CHAIN 22 538 ENVELOPE PROTEIN.			
SO	SEQUENCE 538 AA; 59866 MW; C54648A3C7043870 CRC64;			
Query Match	100.0%; Score 2871; DB 4; Length 538;			
Best Local Similarity	100.0%; Pred. No. 4.9e-248;			
Matches	538; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1 MALPHYHFLFTVLLPSFTLTAPPCRCMTSSSPYQEFILRMQRPNIDAPSRSLSKGP 60			
DB	1 MALPHYHFLFTVLLPSFTLTAPPCRCMTSSSPYQEFILRMQRPNIDAPSRSLSKGP 60			
OY	61 TTTATHTMRNCYHSATLCMHANTHYWNGKMINPSCPGGLGVTCVMTYFTQGMSDGGV 120			

```

Db      61  TFTAHTMPRNCYHSATLCMHANTHYWTGKMINSCGGGLGVTCWTFYQTGMSDGGV 120
Qy      121  ODAAREKHKEVISQLTRVHGTSPPYKGLDLSKHEFLRTHTRLVSLFNTTLGLHEVSA 180
Db      121  ODAAREKHKEVISQLTRVHGTSPPYKGLDLSKHEFLRTHTRLVSLFNTTLGLHEVSA 180
Qy      181  QNPTNCWICLPNRPVYSIPVPEQNNFSTEINTSVLVGPLYSNLEITHTSMLTCVKF 240
Db      181  QNPTNCWICLPNRPVYSIPVPEQNNFSTEINTSVLVGPLYSNLEITHTSMLTCVKF 240
Qy      241  SNTTYTNSOCIRKWTPTPTQIVCLPSGIFVCGTSAYRCLNGSSESCFLSFLVPPMTIY 300
Db      241  SNTTYTNSOCIRKWTPTPTQIVCLPSGIFVCGTSAYRCLNGSSESCFLSFLVPPMTIY 300
Qy      301  TEDDLVSYSKPRNKRPVLPFVIGAGVLCALGTGIGITTSQFYFKLSOELNGDME 360
Db      301  TEDDLVSYSKPRNKRPVLPFVIGAGVLCALGTGIGITTSQFYFKLSOELNGDME 360
Qy      361  VADSLVTLQDQLNSLAAVVQLNRRALDRLTAERGTCLFLGEECCYVYVNGSIYTERVKE 420
Db      361  VADSLVTLQDQLNSLAAVVQLNRRALDRLTAERGTCLFLGEECCYVYVNGSIYTERVKE 420
Qy      421  IRRIORRAEELNRTGFWGLLSQMPWILPFLGLPAAIILLLFGPCIFNLVNFVSRI 480
Db      421  IRRIORRAEELNRTGFWGLLSQMPWILPFLGLPAAIILLLFGPCIFNLVNFVSRI 480
Qy      481  EAVKLOMEPKMOSKTKTYRRPLDRPASPRSDVNDIKGTPPEEISAAPLLRPNAGSS 538
Db      481  EAVKLOMEPKMOSKTKTYRRPLDRPASPRSDVNDIKGTPPEEISAAPLLRPNAGSS 538

```

RESULT 2

```

ID      09NZG3      PRELIMINARY;      PRT;      538 AA.
AC      09NZG3;
DT      01-OCT-2000 (TReMBLrel. 15, Created)
DT      01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DE      Syncytin precursor (Enverlin).
OS      Homo sapiens (Human)
OC      Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Testis;
RA      Sma M., Lee X., Li X., Veldman G.M., Finerly H., Racle L.,
RA      Lavallie E., Tang X., Edoard P., Howes S., Keith J.C. Jr.,
RA      McCoy J.M.;
RT      "Syncytin is a caprine retroviral envelope protein involved in human
RT      placental morphogenesis.";
RL      Nature 403:785-789(2000).
RN      [2]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Placenta;
RA      Alliel P.M., Perin J.P., Pierig R., Nussbaum J.L., Menard A.,
RA      Rieger F.;
RT      "Endogenous retroviruses and multiple sclerosis. Part 2: HERV-7q and
RT      its env transcripts.";
RL      C. R. Acad. Sci. III, Sci. Vie 312:857-863(1998).
RN      [3]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Placenta;
RX      MEDLINE=21965840; PubMed=11990458;
RA      Alliel P.M., Perin J.P., Goudou D., Bitoun M., Robert B., Rieger F.;
RT      "The HERV-W/7q family in the human genome. Potential for protein
RT      expression and gene regulation.";
RL      Cell. Mol. Biol. 48:213-217(2002).
DR      EMBL: AF208161; AAF28334.1; -.
DR      EMBL: AF513360; AAM4759.1; -.
DR      HSSP: P03385; IMOP.

```

```

DR      InterPro: IPR002050; Env_Polyprotein.
DR      Pfam: PF00429; Env_Polyprotein: 1.
KW      Signal.
FT      SIGNAL.
FT      CHAIN
SQ      SEQUENCE 538 AA; 59893 MM; B9498850CB34FB78 CRC64;
Query Match 99.98; Score 2868; DB 4; Length 538;
Best Local Similarity 99.88; Pred. No. 9,1e-248;
Matches 537; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1  MALPYHFLFTVLLPSTLTAPPCRCMTSSSPYOEFLMRQRGNDAPSRSLSGTP 60
Db      1  MALPYHFLFTVLLPSTLTAPPCRCMTSSSPYOEFLMRQRGNDAPSRSLSGTP 60
Qy      61  TFTAHTMPRNCYHSATLCMHANTHYWTGKMINSCGGGLGVTCWTFYQTGMSDGGV 120
Db      61  TFTAHTMPRNCYHSATLCMHANTHYWTGKMINSCGGGLGVTCWTFYQTGMSDGGV 120
Qy      121  ODAAREKHKEVISQLTRVHGTSPPYKGLDLSKHEFLRTHTRLVSLFNTTLGLHEVSA 180
Db      121  ODAAREKHKEVISQLTRVHGTSPPYKGLDLSKHEFLRTHTRLVSLFNTTLGLHEVSA 180
Qy      181  QNPTNCWICLPNRPVYSIPVPEQNNFSTEINTSVLVGPLYSNLEITHTSMLTCVKF 240
Db      181  QNPTNCWICLPNRPVYSIPVPEQNNFSTEINTSVLVGPLYSNLEITHTSMLTCVKF 240
Qy      241  SNTTYTNSOCIRKWTPTPTQIVCLPSGIFVCGTSAYRCLNGSSESCFLSFLVPPMTIY 300
Db      241  SNTTYTNSOCIRKWTPTPTQIVCLPSGIFVCGTSAYRCLNGSSESCFLSFLVPPMTIY 300
Qy      301  TEDDLVSYSKPRNKRPVLPFVIGAGVLCALGTGIGITTSQFYFKLSOELNGDME 360
Db      301  TEDDLVSYSKPRNKRPVLPFVIGAGVLCALGTGIGITTSQFYFKLSOELNGDME 360
Qy      361  VADSLVTLQDQLNSLAAVVQLNRRALDRLTAERGTCLFLGEECCYVYVNGSIYTERVKE 420
Db      361  VADSLVTLQDQLNSLAAVVQLNRRALDRLTAERGTCLFLGEECCYVYVNGSIYTERVKE 420
Qy      421  IRRIORRAEELNRTGFWGLLSQMPWILPFLGLPAAIILLLFGPCIFNLVNFVSRI 480
Db      421  IRRIORRAEELNRTGFWGLLSQMPWILPFLGLPAAIILLLFGPCIFNLVNFVSRI 480
Qy      481  EAVKLOMEPKMOSKTKTYRRPLDRPASPRSDVNDIKGTPPEEISAAPLLRPNAGSS 538
Db      481  EAVKLOMEPKMOSKTKTYRRPLDRPASPRSDVNDIKGTPPEEISAAPLLRPNAGSS 538

```

RESULT 3

```

ID      09NR22      PRELIMINARY;      PRT;      538 AA.
AC      09NR22;
DT      01-OCT-2000 (TReMBLrel. 15, Created)
DT      01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DE      Envelope protein.
OS      Homo sapiens (Human)
OC      Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=20284713; PubMed=10826480;
RA      Voisset C., Bouton O., Bedin F., Duret L., Mandrand B., Maillet F.,
RA      Paranhos-Baccala G.;
RT      "Chromosomal distribution and coding capacity of the human endogenous
RT      retrovirus HERV-W family.";
RL      AIDS Res. Hum. Retroviruses 16:731-740(2000).
DR      EMBL: AF156963; AAF74215.1; -.
DR      HSSP: P03385; IMOP.
DR      InterPro: IPR002050; Env_Polyprotein.

```


Query Match 86.7%; Score 2490; DB 15; Length 542;
 Best Local Similarity 87.3%; Pred. No. 6,1e-214;
 Matches 473; Conservative 17; Mismatches 48; Indels 4; Gaps 1;

```

QY 1 MALPYHIFLFTVLLPSTLTAPPCRCMTSSPYOEFLMRQRCGNIDAPSYRLSKGTP 60
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1 MALPYHIFLFTVLLPSTLTAPPCRCMTSSPYOEFLMRQRCGNIDAPSYRLSKGNS 60
QY 61 TFTAHTMPRNCYSATLCMHANTHYMTGKMINSCGCGAGVTCWTFYFOTGSDGGV 120
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 61 TFTAHTMPRNCYSATLCMHANTHYMTGKMINSCGCGAGVTCWTFYFOTGSDGGI 120
QY 121 QDQAREKHVKEVISQLTRVHGTSPPYKGLDLSKHETLRHTRLVSLFNTTLTGHEVSA 180
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 121 QDQAREKHVKEVISQLTRVHGTSPPYKGLDLSKHETLRHTRLVSLFNTTLTGHEVSA 180
QY 181 QNPNCWICLPANRPVYSIPVPEQMNNSFEITNTSVLVGPLYNSLEITHSNLTCYKF 240
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 181 QNPNCWICLPANRPVYSIPVPEQMNNSFEITNTSVLVGPLYNSLEITHSNLTCYKF 240
QY 241 SNTYTTNSOCIRWVTPPTQIVCLPSGIFVCGTSAYRCNLGSSSCMCFSLFVPPMTY 300
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 241 SNTYTTNSOCIRWVTPPTQIVCLPSGIFVCGTSAYRCNLGSSSCMCFSLFVPPMTY 300
QY 301 TEODLYSVISKPRNKRVPILPFYIGAGVGLGALGTGIGITTSQFYFKLSQELNGDME 360
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 301 TEODLYSVISKPRNKRVPILPFYIGAGVGLGALGTGIGITTSQFYFKLSQELNGDME 360
QY 361 VADSLVTLQDQNSLAVALVIONRRALDLTAERCGTCLFGEBCCYVYVNO 420
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 361 VADSLVTLQDQNSLAVALVIONRRALDLTAERCGTCLFGEBCCYVYVNO 420
QY 421 IRDRIORARELRTGPMGLLSOMPMWLPPLGFLAIIILLLFGPCIFNLVNFVSSRI 480
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 421 IRDRIORARELRTGPMGLLSOMPMWLPPLGFLAIIILLLFGPCIFNLVNFVSSRI 480
QY 481 EAVK----LQMEPKQSKTKIYRRPLDRPASPRSDVNDIKGTPEETSAAQPLRPVSAG 536
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 481 EAVK----LQMEPKQSKTKIYRRPLDRPASPRSDVNDIKGTPEETSAAQPLRPVSAG 536
QY 537 SS 538
   ||
DB 541 SS 542

```

RESULT 6

```

ID 095244 PRELIMINARY; PRT; 410 AA.
AC 095244;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Envelope protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=9909005; PubMed=9882319;
RA Blond J.L., Beseme F., Duret L., Bouton O., Bedin F., Perron H.,
RT Mandrand B., Mallet F.;
RT "Molecular characterization and placental expression of HERV-W, a new
RT human endogenous retrovirus family.";
RL EMBL: AF072505; AAD14545.1; -.
DR HSSP: P03385; IMOF.
DR InterPro: IPR002050; Env_poli_protein.
DR Pfam: PF00429; Env_poli_protein; 1.
FT NON_TER 410
SQ SEQUENCE 410 AA; 45442 MW; 6371EC35DE1587 CRC64;

```

Query Match 76.4%; Score 2194; DB 4; Length 410;
 Best Local Similarity 99.0%; Pred. No. 1,3e-187;
 Matches 406; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```

QY 1 MALPYHIFLFTVLLPSTLTAPPCRCMTSSPYOEFLMRQRCGNIDAPSYRLSKGTP 60
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1 MALPYHIFLFTVLLPSTLTAPPCRCMTSSPYOEFLMRQRCGNIDAPSYRLSKGTP 60
QY 61 TFTAHTMPRNCYSATLCMHANTHYMTGKMINSCGCGAGVTCWTFYFOTGSDGGV 120
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 61 TFTAHTMPRNCYSATLCMHANTHYMTGKMINSCGCGAGVTCWTFYFOTGSDGGV 120
QY 121 QDQAREKHVKEVISQLTRVHGTSPPYKGLDLSKHETLRHTRLVSLFNTTLTGHEVSA 180
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 121 QDQAREKHVKEVISQLTRVHGTSPPYKGLDLSKHETLRHTRLVSLFNTTLTGHEVSA 180
QY 181 QNPNCWICLPANRPVYSIPVPEQMNNSFEITNTSVLVGPLYNSLEITHSNLTCYKF 240
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 181 QNPNCWICLPANRPVYSIPVPEQMNNSFEITNTSVLVGPLYNSLEITHSNLTCYKF 240
QY 241 SNTYTTNSOCIRWVTPPTQIVCLPSGIFVCGTSAYRCNLGSSSCMCFSLFVPPMTY 300
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 241 SNTYTTNSOCIRWVTPPTQIVCLPSGIFVCGTSAYRCNLGSSSCMCFSLFVPPMTY 300
QY 301 TEODLYSVISKPRNKRVPILPFYIGAGVGLGALGTGIGITTSQFYFKLSQELNGDME 360
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 301 TEODLYSVISKPRNKRVPILPFYIGAGVGLGALGTGIGITTSQFYFKLSQELNGDME 360
QY 361 VADSLVTLQDQNSLAVALVIONRRALDLTAERCGTCLFGEBCCYVYVNO 410
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 361 VADSLVTLQDQNSLAVALVIONRRALDLTAERCGTCLFGEBCCYVYVNO 410

```

RESULT 7

```

ID 096159 PRELIMINARY; PRT; 319 AA.
AC 096159;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Envelope protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Jern P., Karlsson D., Blomberg J.;
RT "New HERV-W SU sequences.";
RL Submitted (AUG-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL: AY050301; AAL11495.1; -.
FT NON_TER 319
SQ SEQUENCE 319 AA; 35784 MW; 7D288CCCD1B68999 CRC64;

```

Query Match 57.9%; Score 1663; DB 4; Length 319;
 Best Local Similarity 95.0%; Pred. No. 2,9e-140;
 Matches 303; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

```

QY 1 MALPYHIFLFTVLLPSTLTAPPCRCMTSSPYOEFLMRQRCGNIDAPSYRLSKGTP 60
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1 MALPYHIFLFTVLLPSTLTAPPCRCMTSSPYOEFLMRQRCGNIDAPSYRLSKGTP 60
QY 61 TFTAHTMPRNCYSATLCMHANTHYMTGKMINSCGCGAGVTCWTFYFOTGSDGGV 120
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 61 TFTAHTMPRNCYSATLCMHANTHYMTGKMINSCGCGAGVTCWTFYFOTGSDGGV 120
QY 121 QDQAREKHVKEVISQLTRVHGTSPPYKGLDLSKHETLRHTRLVSLFNTTLTGHEVSA 180
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 121 QDQAREKHVKEVISQLTRVHGTSPPYKGLDLSKHETLRHTRLVSLFNTTLTGHEVSA 180
QY 181 QNPNCWICLPANRPVYSIPVPEQMNNSFEITNTSVLVGPLYNSLEITHSNLTCYKF 240
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```


DB 121 AGMLGLGTGMSITTSQFYFKLSOELNGDMERVADSLVTLDQDLSLAAYVLQNRRL 180

RESULT 14

Q96TB3 PRELIMINARY: PRT: 179 AA.
AC Q96TB3;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Envelope protein (Fragment).
GN ENV.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21268768; PubMed=11375061;
RA Kim H.-S., Lee W.-H.;
RT "Human endogenous retrovirus HERV-W family: chromosomal localization, identification, and phylogeny.";
RL AIDS Res. Hum. Retroviruses 17:643-648(2001).
DR EMBL; AB051008; BAB47560.1; -
DR InterPro; IPR002050; Env_polypotein.
DR Pfam; PF00429; ENV_polypotein; 1.
FT NON_TER 1
FT 179
SQ SEQUENCE 179 AA; 19471 MW; 70DF3E5658B8667 CRC64;

Query Match 28.9%; Score 829.5; DB 4; Length 179;
Best Local Similarity 90.6%; Pred. No. 4,8e-66;
Matches 163; Conservative 6; Mismatches 10; Indels 1; Gaps 1;

QY 207 NNFSTEINTSVLVGVLVSNLEITHRSNLTVCVKSNTTYTNSQCIRWTPPTQYVCLPS 266
DB 1 NNFSTEINTSVLVGVLVSNLEITHRSNLTVCVKSNTTYTNSQCIRWTPPTQYVCLPS 59
QY 267 GIFFVCGTSAYRCQLNGSSSEMCFLSFLVPEPMITTEODLXSVISKPRKRVPILTFVYG 326
DB 60 GIFFVCGTSAYRCQLNGSSSEMCFLSFLVPEPMITTEODLXSVISKPRKRVPILTFVYG 119
QY 327 AGVLGALGTGIGITTSQFYFKLSOELNGDMERVADSLVTLDQDLSLAAYVLQNRRL 386
DB 120 AGVLGALGTGIGITTSQFYFKLSOELNGDMERVADSLVTLDQDLSLAAYVLQNRRL 179

RESULT 15

Q96TB2 PRELIMINARY: PRT: 180 AA.
AC Q96TB2;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Envelope protein (Fragment).
GN ENV.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21268768; PubMed=11375061;
RA Kim H.-S., Lee W.-H.;
RT "Human endogenous retrovirus HERV-W family: chromosomal localization, identification, and phylogeny.";
RL AIDS Res. Hum. Retroviruses 17:643-648(2001).
DR EMBL; AB051009; BAB47561.1; -
DR InterPro; IPR002050; Env_polypotein.
DR Pfam; PF00429; ENV_polypotein; 1.
FT NON_TER 1
FT 180
SQ SEQUENCE 180 AA; 19709 MW; 4C8F196C9C788497 CRC64;

QY 207 NNFSTEINTSVLVGVLVSNLEITHRSNLTVCVKSNTTYTNSQCIRWTPPTQYVCLPS 266
DB 1 NNFSTEINTSVLVGVLVSNLEITHRSNLTVCVKSNTTYTNSQCIRWTPPTQYVCLPS 60
QY 267 GIFFVCGTSAYRCQLNGSSSEMCFLSFLVPEPMITTEODLXSVISKPRKRVPILTFVYG 326
DB 61 GIFFVCGTSAYRCQLNGSSSEMCFLSFLVPEPMITTEODLXSVISKPRKRVPILTFVYG 120
QY 327 AGVLGALGTGIGITTSQFYFKLSOELNGDMERVADSLVTLDQDLSLAAYVLQNRRL 386
DB 121 AGVLGALGTGIGITTSQFYFKLSOELNGDMERVADSLVTLDQDLSLAAYVLQNRRL 180

Query Match 28.8%; Score 828; DB 4; Length 180;
Best Local Similarity 88.3%; Pred. No. 6,6e-66;
Matches 159; Conservative 9; Mismatches 12; Indels 0; Gaps 0;
Search completed: September 23, 2003, 14:57:13
Job time : 109 secs

THIS PAGE BLANK (USPTO)